The international journal of science / 15 June 2023

nature

spotlight
Nutrition research in China

CUTTING THROUGH THE NOISE

Error mitigation empowers quantum processor to probe physics that classical methods can't reach

Call of the wild
Tracking natural behaviour in animals to decode the brain

Soda stream
Phosphates found in ice ejected from ocean on Enceladus

Sowing the seeds
Ancient DNA reveals how farming came to northwest Africa

VISIT…

LANZAROTE Caliente.com

Nature.2023.06.17

[Sat, 17 Jun 2023]

- This Week
- News in Focus
- Books & Arts
- Opinion
- Work
- Research
- Amendments & Corrections
- Spotlight

This Week

- **The US EPA's proposed regulation could help to kill off fossil-fuel plants. Good on it** [13 Jun 2023]
  • Expanding clean energy isn't enough to combat the climate crisis — the agency's move to turn the screws on fossil fuels is welcome, and long overdue.
- **Vaccine-derived polio is undermining the fight to eradicate the virus** [14 Jun 2023]
  • Wild polio has almost been eradicated, but vaccine-derived strains retain the potential to paralyse. Better vaccines have arrived — but they are only part of the answer.
- **Tackling pervasive sexism in Australian science requires money, leadership and time** [12 Jun 2023]
  • Lessons in advancing gender equity in research from my time in the United States at Lamont–Doherty Earth Observatory and in Australia.

- EDITORIAL
- 13 June 2023

The US EPA's proposed regulation could help to kill off fossil-fuel plants. Good on it

Expanding clean energy isn't enough to combat the climate crisis — the agency's move to turn the screws on fossil fuels is welcome, and long overdue.



Coal-fired power plant in Utah, one of the states opposing the EPA's proposed emissions rule.Credit: George Frey/AFP/Getty

Last month, operators of the main fossil-fuel power plants in the United States were put on notice: if they want to continue operating after 2040, they

would need to reduce their carbon footprint by at least 90%. In the case of power plants that run on coal, the dirtiest of the fossil fuels, that potentially leaves only one option to avoid closure: these plants must capture and bury their emissions using carbon-capture and storage (CCS) technology.

The proposed rule, announced by the Environmental Protection Agency (EPA), is a belated but welcome step. It is not enough to expand clean-energy technologies: governments must also tackle existing sources of emissions. If implemented, the EPA plan would, albeit slowly, do just that. It would also send a powerful signal that one of the world's largest greenhouse-gas emitters is serious about phasing out conventional fossil-fuel facilities. If anything, the EPA's rule could be strengthened by bringing the end date forward and applying these requirements to more power plants.



Carbon capture key to Biden's new power-plant rule: is the tech ready?

Over little more than a decade, the United States has cut the quantity of electricity it generates using coal by more than half, from 1.7 trillion kilowatt-hours in 2011 to 828 billion in 2022, a trend that is set to continue. But the use of natural gas for electricity generation has been rising steadily in recent years, alongside renewable sources such as wind turbines and solar panels. As it stands, the United States still counts on fossil fuels for around 60% of its electricity — of these fuels, two-thirds is gas and one-third coal. That is why the EPA's rule is so important. By setting such a high bar for

emissions reductions by power plants, it effectively mandates the use of CCS, if the highest-emitting facilities are to continue operating, and that would be a game-changer.

Trial and error

Long-standing research and development efforts suggest that CCS is a viable technology to reduce power-plant emissions — indeed, it has been trialled at numerous pilot plants around the world. Some have encountered snags: equipment problems have frequently limited carbon-capture operations at SaskPower's Boundary Dam, a coal-fired power plant in Saskatchewan, Canada, for example. But researchers say that such technological kinks can be straightened out as operators gain experience. The main reason CCS has never taken off commercially — and has only rarely been operated at full scale — is economic. Unless governments actively step in to mandate CCS or put a sufficiently high price on carbon emissions, it will be cheaper to emit carbon into the atmosphere. The proposed regulation would shift that calculation.

Not surprisingly, a coalition that includes states that rely on fossil-fuel extraction industries is preparing to challenge the EPA in court. There is a precedent: under former president Barack Obama, the EPA had crafted a regulation that sought to broadly shift power generation across the grid towards cleaner forms of energy. However, the Supreme Court ruled last year that the EPA doesn't have the authority to oversee the restructuring of the entire electrical grid, and that its remit is limited to mandating technologies that can be used at individual power plants.



[US Supreme Court hobbles the EPA's authority over climate emissions]

The EPA's current proposal seeks to comply with that ruling by arguing that CCS represents a viable technology that power-plant operators can install to slash the emissions of their facilities. The EPA's challengers are expected to argue that the technology is too costly and unproven. This means that whether the proposal stands up will depend mostly on whether the court agrees that CCS is ready for prime time.

To be clear, CCS is not a panacea for the power sector. Even if power plants are able to capture 90% of their emissions, the remaining 10% will continue to be pumped into the atmosphere. At the same time, the costly and environmentally damaging extraction of fossil fuels would continue. In most cases, it would make more sense to shut down fossil-fuel power plants and transition to truly clean energy.

And that could indeed be the effect. In the case of coal-fired power plants, many will probably close down instead of complying with the proposed regulation, as EPA administrator Michael Regan acknowledges. The same could be true for many large gas-fired power plants, faced with the choice of adopting CCS or a costly conversion to burning 'green' hydrogen to meet the requirements.



How to rebuild the US Environmental Protection Agency

Other regulatory requirements on fossil-fuel burning are also growing, even as the price of renewable-energy generation is tumbling. This means both pain and challenges as communities, states and businesses make decisions about which fossil-fuel plants to close, and where and how to fill the gap with clean energy. Many jobs will be lost, and others created. Those that are created will require new skills and won't necessarily be in the same location.

The administration of President Joe Biden seems to be aware of these social and economic (and political) realities. In April, the White House announced that new policies have already directed more than US$14 billion in federal investments towards communities struggling with the loss of fossil-fuel-related jobs, with more investments to come. Managing the social costs of the clean-energy transition must remain a priority in the United States, and worldwide.

Everyone will all be better off for it. The simple fact is that fossil fuels are dirty, from end to end. The air pollution they create kills millions of people each year around the globe. The greenhouse gases they pump into the atmosphere are driving a climate crisis that is already threatening people and natural ecosystems worldwide. Ultimately, to halt global warming, greenhouse-gas emissions must be eliminated or offset by carbon uptake

elsewhere. This means making difficult choices — and beginning, with intent, to tighten the chokehold on fossil fuels.

Nature **618**, 433 (2023)

doi: https://doi.org/10.1038/d41586-023-01825-0

This article was downloaded by **calibre** from https://www.nature.com/articles/d41586-023-01825-0

- EDITORIAL
- 14 June 2023

Vaccine-derived polio is undermining the fight to eradicate the virus

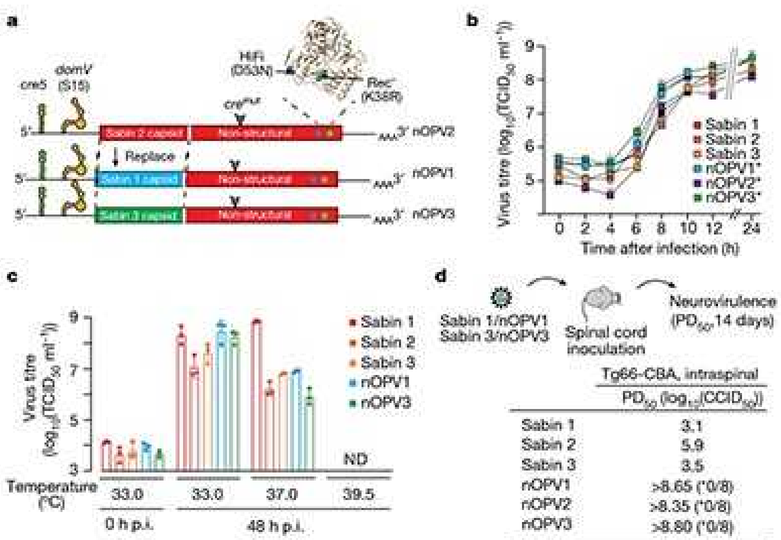
Wild polio has almost been eradicated, but vaccine-derived strains retain the potential to paralyse. Better vaccines have arrived — but they are only part of the answer.



The nOPV2 polio vaccine has been administered to more than 650 million children across 30 countries.Credit: Issouf Sanogo/AFP/Getty

When smallpox was declared eradicated in 1980, there were hopes that poliovirus might soon be next. As a result of an eradication campaign that began in 1988, two of the virus's three subtypes have since been eradicated and the third has been cornered. In the past 12 months, just 21 cases of wild polio were reported in children across Afghanistan, Pakistan and Mozambique.

But cases of another type of polio are worryingly high. Over the past year, there have been 665 cases of vaccine-derived polio across 23 countries. Vaccine-derived polio occurs when the attenuated poliovirus used to make oral vaccines regains its virulence.



[Read the paper: Genetic stabilization of attenuated oral vaccines against poliovirus types 1 and 3]

In *Nature* this week, Andrew Macadam at the National Institute for Biological Standards and Control in South Mimms, UK, and his colleagues report the development of novel oral polio vaccines that reduce the likelihood of vaccine-derived outbreaks occurring for poliovirus types 1 and 3[1]. These vaccines were created using an approach [previously used to make a new type 2 polio vaccine] known as nOPV2[2]. The scientists used genetic engineering to slash the likelihood that the attenuated virus will revert to virulence. The two latest vaccines were tested in mice and found to be effective, safe and stable. They are now being tested in human trials. If their

safety and efficacy are as good as that of the nOPV2 vaccine, all three poliovirus subtypes will be covered.

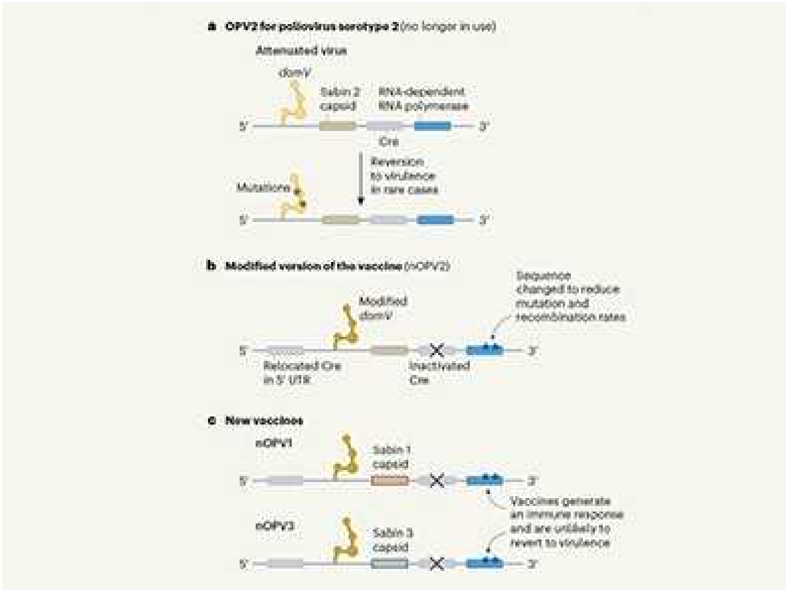
Since March 2021, the nOPV2 vaccine has been administered to more than 650 million children across 30 countries. But although nOPV2 is much more stable genetically than its predecessor, it can regain virulence, too. So far, it has done so on four separate occasions, in Nigeria and the Democratic Republic of the Congo. This is probably because eradication campaigns are not reaching enough children, and it reinforces what researchers and public-health officials already know: for polio to be eradicated, advances in science need to go hand-in-hand with comprehensive public-health initiatives.

Why cases are going up

Children in under-immunized communities are especially at risk of outbreaks of vaccine-derived polio. In the past few years, most outbreaks have occurred in sub-Saharan Africa. However, vaccine-derived polio has also been identified in sewage samples in the United Kingdom, Canada, Israel and the United States — and the last two countries have each seen a case of paralytic vaccine-derived polio.

The rise in the number of outbreaks of vaccine-derived polio has various causes. The COVID-19 pandemic hindered vaccination campaigns and surveillance. Disasters and conflicts, which often occur in countries vulnerable to polio, have done the same. Moreover, some countries decided to delay their outbreak responses and wait for supplies of the nOPV2 vaccine, rather than use the older type 2 vaccine. A modelling study, originally published in 2021, suggested that this delay would increase the risk of outbreaks occurring[3].

High-level changes to the way in which polio campaigns are run is another factor underlying the deterioration of some campaigns, according to several researchers that *Nature* spoke to for this editorial.



[Polio endgame finish is in sight]

Around the world, polio eradication has long been handled by the Global Polio Eradication Initiative (GPEI), a partnership based in Geneva, Switzerland, that includes national governments, philanthropic funders and the World Health Organization. However, the WHO is now overseeing a plan to transfer some of the GPEI's activities to national immunization and disease-surveillance programmes. This has involved the GPEI handing over some responsibility to the WHO, and the WHO then helping national programmes to manage independently.

The transition has not been easy. There have been funding gaps, expert staff have been lost, governance has faltered in some countries and morale has dropped in others. A report published in April 2022 by the Independent Monitoring Board (IMB), a group of global-health experts that the GPEI commissions to produce periodic assessments of progress, has been critical of the transition process in some countries (see go.nature.com/3qhjk6a).

A WHO spokesperson told *Nature* that, after commissioning a more recent independent report, the transition programme accepts that it "pulled the plug too early in some of these countries", and said that it is drawing up a more nuanced, country-by-country approach.

The publication of a new independent report later this month will reveal more about how the transition is going. In the long term, it makes sense to integrate the GPEI's formidable polio resources into national immunization programmes. But it is crucial that this be done in a way that does not undermine the progress already made. The GPEI needs to remain in the picture.



Spate of polio outbreaks worldwide puts scientists on alert

Even the use of the word 'transition', combined with the knowledge that wild poliovirus is gone — or almost gone — can, unfortunately, have a dampening effect on the commitment to carry out polio campaigns, as the IMB report points out. In Nigeria in particular, the announcement in 2020 that wild poliovirus had been eliminated created a complacency ill-suited to tackling vaccine-derived polio.

Meanwhile, scientists have other tools in the pipeline to further the polio-eradication campaign. At present, it can take many weeks to confirm that a wastewater sample contains poliovirus. Several research groups are working on technologies to reduce this time lag.

Separately, Macadam and his colleagues are developing a form of the virus that is inactive *in vivo*. They hope this could one day replace the infectious virus that is currently used to make vaccines. And scientists led by David

Rowlands, who is based at the University of Leeds, UK, are developing a vaccine made from virus-like particles. If successful, this would mean no live virus would be needed to make vaccines.

It's essential that researchers continue to advance vaccine research and technology to combat the rise of vaccine-derived polio. But new vaccines will not work if they remain in their vials. If the world is to see the back of vaccine-derived polio, the fruits of vaccine research and development need to be put into the hands of highly motivated, well-funded staff operating in programmes attuned to the local landscape, and the transition of polio management from the GPEI to national control must not undermine eradication. Only then might we be able to send polio the same way as smallpox.

Nature **618**, 434 (2023)

doi: https://doi.org/10.1038/d41586-023-01953-7

References

1. Yeh, M. T. *et al. Nature* https://doi.org/10.1038/s41586-023-06212-3 (2023).
2. Yeh, M. T. *et al. Cell Host Microbe* **27**, 736–751 (2020).
3. Kalkowska, D. A., Pallansch, M. A., Wassilak, S. G. F., Cochi, S. L. & Thompson, K. M. *Vaccine* **41** (Suppl. 1), A136–A141 (2023).

This article was downloaded by **calibre** from https://www.nature.com/articles/d41586-023-01953-7

- WORLD VIEW
- 12 June 2023

Tackling pervasive sexism in Australian science requires money, leadership and time



Lessons in advancing gender equity in research from my time in the United States at Lamont–Doherty Earth Observatory and in Australia.

- Robin Bell

I'm no stranger to tackling sexism in science. From 2004 to 2010, I led a programme to diversify the Lamont–Doherty Earth Observatory of Columbia University in Palisades, New York, by changing its culture. When I started, only two members of Lamont's teaching faculty were women — less than 10%. Today, more than half of faculty members are women.

For the past five months, I've been a visiting scientist as the Fulbright distinguished chair in science, technology and innovation at the Commonwealth Scientific and Industrial Research Organisation (CSIRO), Australia's national science agency in Hobart. One of my goals was to share evidence-based practices to diversify scientific workplaces.

In April, a scathing report revealed gender inequity, bullying and sexual harassment in the Australian Antarctic Division. Three days after its release, tears filled people's eyes during my talk on fostering culture change in polar research when I spoke of the need to fix the system, not the women.



To advance equality for women, use the evidence

To advance gender equity in science, Australia — like every country — needs a broad, well-funded federal programme that implements evidence-based approaches, supports the selection of scientific leaders committed to diversity and builds in accountability. I can't say exactly what the solution will be because it must be grounded in the realities of the Australian system, as ours was at Lamont. I am sure it will take time, so long-term funding and patience will be crucial.

Female scientists in Australia have shared stories of bullying, harassment, bias and unwelcoming cultures, of being passed over for leadership roles, of having ideas stolen or being told that sexism is just part of the process.

Male leaders tell stories of their paths to success through old boys' networks without acknowledging that such a path is inaccessible to many. According to a [2019 report by the Champions of Change Coalition](), an Australia-based group of leaders advancing gender equity, 66% of women in Australian science have had their voices devalued. [Nearly half the respondents to another 2019 survey have been sexually harassed]().

I'm an outsider — I don't know more about Australian science than do those who have been there for years. What I do know is what worked at Lamont, and it wasn't necessarily what we expected.

First, having money gave us credibility. The US National Science Foundation has invested more than US$270 million through its ADVANCE programme, targeting institutional change. This financial commitment has changed the culture at more than 170 institutions, including Lamont, and produced a large cadre of diversity champions, a body of knowledge on institutional change and leaders versed in evidence-based practices.



[Scientists say harassment in the Antarctic must stop — but US plan falls short]()

Second, it took time. This work is difficult and slow. Institutional change was harder than running any Antarctic field programme I've ever led. After the first five years I thought we had failed, but, after another five, we had

placed women in academic leadership positions all over the country. Implementing a visiting fellowship for women has helped us hire prominent female scholars. Maureen Raymo became Lamont's first female director in 2020.

Third, each institution needs to identify its specific problems and design its own equitable future, empowered by the knowledge that diversity matters. Before we analysed our system, I was convinced that childcare during fieldwork was one of our biggest stumbling blocks for gender equity because of the trouble I had had earlier in my career, but fieldwork wasn't one of the most important issues. We found that making sure that women attended meetings and launched workshops — so the community could see them as intellectual leaders — was more important.

Some Australian organizations have been putting in the hard work, time and money necessary to shift the needle on gender equity. For example, in 2013, none of the senior leaders of the government agency Geoscience Australia was a woman, and it feared that its rising female stars would leave if they could not see themselves thriving if the environment was perceived as hostile and male-dominated. Insights from a cultural audit allowed Geoscience Australia to confront embedded biases and get all of its staff members to define what an inclusive workplace would look like. Now, nearly half the senior leaders are female. Diversity is not a side note but is crucial to everything in the agency's strategic plan.

Such efforts to change the culture should be replicated and robustly funded elsewhere. This will require leadership, money and patience. Change efforts should be connected, not fragmented, so that people can support and learn from each other, and accountability should have teeth. Australian national-science leaders need to speak up about how important an inclusive culture is to retain research talent.

A place to start implementing a well-funded federal programme could be the current review by the Australian Department of Industry, Science and Resources of women in science, technology, engineering and mathematics programmes. This began in October 2022 and draft recommendations are due in July, with final recommendations expected in October. Another hopeful sign is that the recently passed Respect@Work legislation requires

employers to be proactive in preventing to prevent sexual harassment and discrimination.

Through investment and accountability, any country can make science a place where women can thrive.

Nature **618**, 435 (2023)

doi: https://doi.org/10.1038/d41586-023-01925-x

This article was downloaded by **calibre** from https://www.nature.com/articles/d41586-023-01925-x

News in Focus

- **Why are the Canadian wildfires so bad this year?** [09 Jun 2023]
  • Hot, dry weather and human carelessness have led to a huge burnt area — and to a choking haze that is affecting millions of people.
- **Ukraine dam collapse: what scientists are watching** [09 Jun 2023]
  • Extensive flooding could have severe consequences for farming, health and the environment.
- **Morality is declining, right? Scientists say that idea is an illusion** [07 Jun 2023]
  • Surveys show people around the world have believed for decades that morals are decaying — but other survey data contradict that perception.
- **China's rolling COVID waves could hit every six months — infecting millions** [07 Jun 2023]
  • The latest surge is unlikely to crash the country's health-care system, but scientists fear hundreds of millions of infections.
- **DeepMind AI creates algorithms that sort data faster than those built by people** [07 Jun 2023]
  • The technology developed by DeepMind that plays Go and chess can also help to write code.
- **Colombia's 'cocaine hippo' population is even bigger than scientists thought** [02 Jun 2023]
  • The most comprehensive census yet reveals that there could be twice as many of the invasive animals than previous estimates indicated.
- **'Science was heard': woman who was convicted of killing her children pardoned after inquiry** [06 Jun 2023]
  • Kathleen Folbigg, who was jailed in Australia in 2003 over the sudden deaths of her four young children, has been pardoned and released in the wake of new scientific evidence.
- **Tanzania's researchers offered US$22,000 to publish in international journals** [06 Jun 2023]
  • The move is designed to encourage research and help boost institutions' rankings. But some researchers say it will reward those already established in their careers.

- **[Lab mice go wild: making experiments more natural in order to decode the brain]** [14 Jun 2023]
  • Armed with technologies to track a creature’s every move, neuroscientists are gaining insights into animal — and human — behaviour.
- **[‘It’s a vote for hope’: first gene therapy for muscular dystrophy nears approval, but will it work?]** [02 Jun 2023]
  • The FDA’s decision, expected this month, follows several setbacks and delays and will pose difficult choices for the families of children with Duchenne muscular dystrophy.

| Next section | Main menu | Previous section |

- NEWS
- 09 June 2023

Why are the Canadian wildfires so bad this year?

Hot, dry weather and human carelessness have led to a huge burnt area — and to a choking haze that is affecting millions of people.

- Brian Owens



Smoke billows from a wildfire near Port Alberni in western Canada on 6 June.Credit: James MacDonald/Bloomberg via Getty

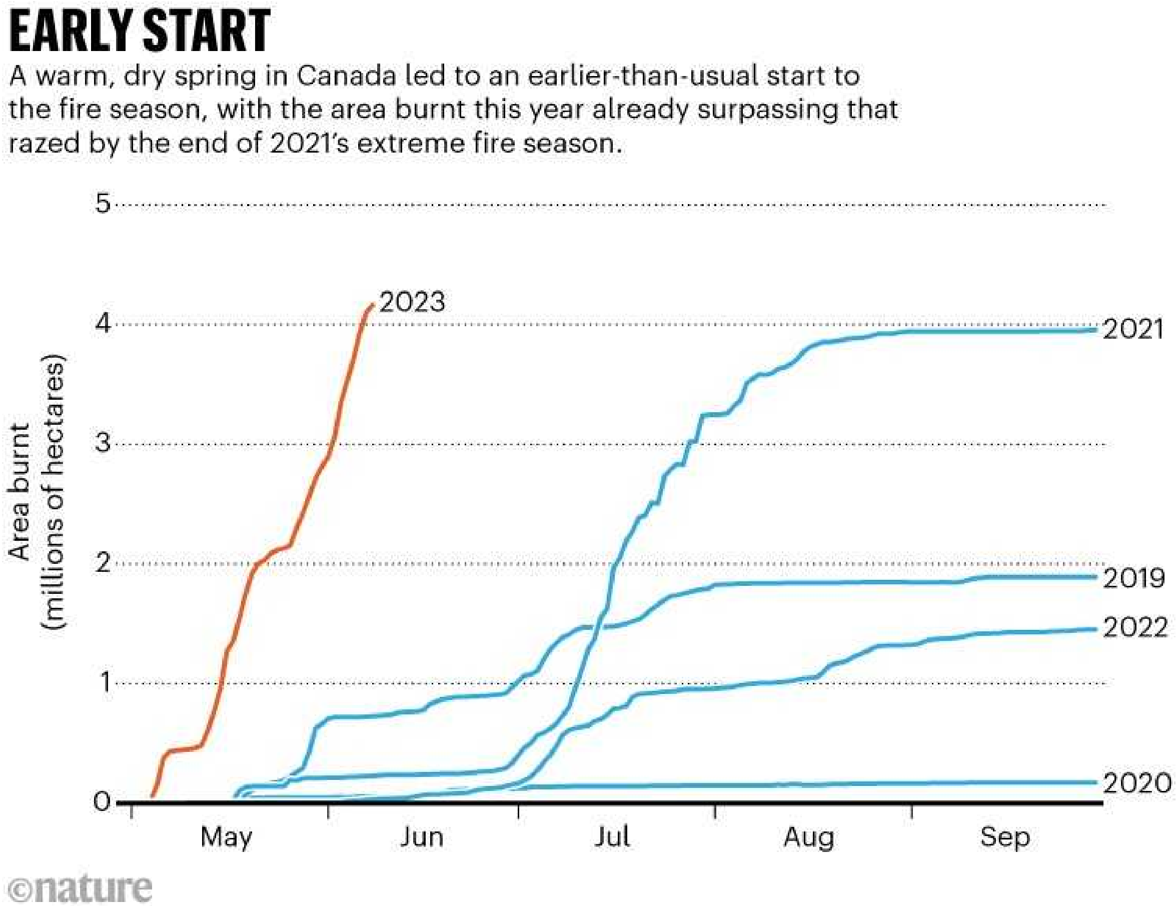
Smoke from wildfires raging in eastern Canada has been filling lungs and turning skies orange across the northeastern United States, most dramatically in New York City and the surrounding area, for the past week. Although people in western North America have become used to such conditions in the past few years, the events are unusual in the eastern part of the continent — so what is causing this extraordinary fire season?

"The number one cause is the weather," says Anthony Taylor, a forest-management specialist at the University of New Brunswick in Fredericton. Fires in Canada are routine, but it has been an especially warm and dry spring across much of the country. That's particularly true in eastern Canada, which had around 50% less spring precipitation than usual, he says. In western Canada, this May was the warmest and driest on record, says David Phillips, a climatologist at Environment and Climate Change Canada.

No blaming El Niño

There is no solid explanation for this year's anomalous spring. Piyush Jain, a research scientist at the Canadian Forest Service, says it is unlikely to be related to an El Niño climate pattern, which brings warmer temperatures to the eastern Pacific Ocean and tends to warm the planet as a whole. The US National Oceanographic and Atmospheric Administration confirmed yesterday that an El Niño has arrived, but the effects are not expected to show up until later in the year. Still, extraordinary weather is not unexpected as the planet warms. "Climate change is definitely a factor that is causing these extreme conditions to occur more frequently," says Jain.

The warmth and lack of rain has left soils and forests as dry as tinder, so when a fire ignites it can grow and spread quickly — including in places where large, destructive fires are usually rare, such as the eastern provinces of New Brunswick and Nova Scotia. More than 4 million hectares of forest have already burnt across Canada this year — double the historical average and a number that is usually reached much later in the season (see 'Early start').



Source: Canadian Wildland Fire Information System

Nor is it common for the smoke to affect so many people in the eastern United States. The reason for the orange skies in New York City and elsewhere this week is a large low-pressure system that has been sitting over Maine for several days. The system is blocking transport of the smoke to the east, while the system's counter-clockwise winds are acting like a conveyor belt, dragging the smoke south to the eastern seaboard, says Jain. The system should start to break up over the weekend, bringing some relief.

The on-the-ground causes of the fires vary across the country. Fires in the broad band burning across the middle of the province of Quebec, whose smoke is travelling to New York City and other US sites, were probably ignited by lightning, says Marc-André Parisien, a research scientist at the Canadian Forest Service. Lightning is usually responsible for about half of all fires in Canada, and 85% of the area burnt each year.

Human mistakes

But this year, most of the fires in the western provinces and on the east coast came earlier in the year than lightning is typically reported and were probably caused by human activity. One fire near St Andrews, in New Brunswick province, for example, started when an all-terrain vehicle caught fire on a trail, igniting the surrounding woods. “The biggest thing we need to do, especially in the east, is make people aware of their role in fires,” says Taylor.

The year’s extreme fire season is part of an ongoing trend in Canada and around the world, says Parisien. Canada has more large fires, they are burning larger areas and the fire season is getting longer — it now starts about a week earlier and ends a week later than it did 50 years ago[1]. Extreme weather that promotes fire — hot, dry and windy — is becoming more common both in Canada and globally[2].

Despite the huge areas that have already burnt this year, we are not out of the woods yet. Depending on the weather, the fires could continue for months with little respite. “If the warm, dry weather continues, there is lots of fuel in the forest to burn,” says Taylor. “It’s not going to be exhausted any time soon.”

Nature **618**, 439-440 (2023)

doi: https://doi.org/10.1038/d41586-023-01902-4

References

1. Hanes, C. C. *et al. Can. J. Forest Res.* **49**, 256–269 (2018).

2. Jain, P., Wang, X. & Flannigan, M. D. *Int. J. Wildland Fire* **26**, 1009–1020 (2017).

This article was downloaded by **calibre** from https://www.nature.com/articles/d41586-023-01902-4

- NEWS EXPLAINER
- 09 June 2023

Ukraine dam collapse: what scientists are watching

Extensive flooding could have severe consequences for farming, health and the environment.

- Miryam Naddaf



Large sections of the Kakhovka dam have collapsed, unleashing catastrophic floods. Credit: Satellite image (c) 2023 Maxar Technologies via Getty

The 66-year-old Kakhovka dam on the Dnieper River in south Ukraine collapsed on the morning of 6 June after a suspected explosion, triggering a

catastrophic humanitarian and environmental crisis.

Spanning an area of more than 2,000 square kilometres, the dam's reservoir is the country's largest in terms of water volume. The dam has been controlled by Russian forces for more than a year.

The breach triggered extensive flooding, which peaked at a depth of 5.6 metres in Kherson on 8 June and has already displaced more than 20,000 people across dozens of settlements, including in Russian-held areas on the river's lower-left bank. The deluge is expected to continue for at least a week.

Nature spoke to researchers and specialists on rivers and environmental science about the continuing impacts of the disaster.

What are some of the immediate consequences?

Before the breach, the Kakhovka reservoir held more than 19 cubic kilometres of water. "Now, there are only 11 cubic kilometres of water left," said Oleksandr Krasnolutskyi, Ukraine's deputy minister of environmental protection and natural resources, in Kyiv, at a press briefing on 8 June.

The reservoir provides water for more than 700,000 people in south Ukraine. Cities on the Dnieper River, including Kherson, Nikopol, Marhanets and Pokrov, are short of water supplies, according to the United Nations.

And the flood waters themselves have caused extensive damage — destroying houses, roads and other crucial infrastructure.

How might the flooding affect farming and food security?

As the water level continues to drop, there will not be enough water for the irrigation canals that the reservoir usually serves, says Roger Falconer, a water engineer at Cardiff University, UK, who specializes in modelling dam failures. "It could affect crops both downstream and upstream."

Flood water has inundated large areas of farms and arable lands, washing away their topsoil layers, according to Ukraine's environment ministry. "We will not be able to cultivate agricultural plants on this soil for many years," said Krasnolutskyi. Falconer adds that the floods could wash fertilizers used on agricultural land into the river, where they could disrupt aquatic ecosystems.

What are the other environmental impacts?

The sudden surge of water downstream has had immediate and far-reaching impacts on the biodiverse ecosystems. "Nearly 160,000 animals and 20,000 birds are under threat because of the catastrophe," said Krasnolutskyi.

Some of those species are rare, or found only in this area. These include the vulnerable Nordmann's birch mouse (*Sicista loriger*) and the endangered sand mole rat (*Spalax arenarius*), according to a report by the Ukrainian Nature Conservation Group (UNCG) in Vasylkiv.

The Kakhovka reservoir itself is home to dozens of fish species. The rapid draining of its water means that vast numbers of fish will be either stranded in shallow, dried-up zones, or swept away to sea, where they will perish in the salt water.

"What we have seen is the tip of the iceberg," says Oleksii Vasyliuk, an environmentalist and co-founder of UNCG. "This is ecocide."

Nearby national parks have also been flooded, which will cause irreparable damage to their flora and fauna.

Nine sites in Ukraine's Emerald Network, a Europe-wide conserved area, as well as five internationally important wetlands have been flooded. Around 55,000 hectares of forest have been inundated with water that is predicted to remain stagnant for 20 days, according to the environment ministry.



A member of Ukraine’s National Guard brings food to Kherson residents stranded by flooding.Credit: Genya Savilov/AFP via Getty

Does the dam’s proximity to a nuclear power plant pose a danger?

Europe’s largest nuclear power plant, in Zaporizhzhia, is located around 150 kilometres upstream of the Kakhovka dam. The plant’s six reactors have been shut down for more than eight months — but it needs cooling water to manage the residual decay heat. The reactor is continuing to pump cooling water in from the reservoir, according to the International Atomic Energy Agency.

If the water level in the Kakhovka reservoir drops too low to be able to supply cooling water, Zaporizhzhia can switch to alternative water supplies. There are also two cooling towers that use the atmosphere for cooling, and require only a small amount of water to operate, says Malte Jansen, an energy scientist at the University of Sussex in Brighton, UK.

Perhaps more concerning is the potential dispersal of toxic compounds. More than 150 tonnes of machine oil from the Kakhovka hydroelectric power station, which sits on the dam, have spilled into the Dnieper River, according to the environment ministry. The flood water also carried garbage, together with construction waste and sewage, into the Dnieper watershed, according to Krasnolutskyi, where it could potentially contaminate supplies of drinking water.

What can be done to address the situation?

If the reservoir's water level continues to fall, it will eventually return to the baseline level before the dam was built, says Falconer.

He adds that the collapse will ultimately change the reservoir's bed topography, and that it could also increase shear stress, the force of water flowing against the river bed, which would disturb any toxic sediment there.

"It will be necessary to either plant [upstream] areas with a forest, or sow meadow grasses so that the wind does not blow away this silt at the bottom of the dried reservoir, because it is contaminated with waste from Zaporizhzhia," says Vasyliuk.

The environment ministry says that a scientific survey will be needed to explore whether the dam should be rebuilt. But a complete assessment of the flood's impact is unlikely at present: Russian forces currently control the south side of the river, where most of the flooding has occurred. "Nothing can be done to minimize the consequences," says Vasyliuk. "This is a zone of both an environmental disaster and active hostilities."

Nature **618**, 440-441 (2023)

doi: https://doi.org/10.1038/d41586-023-01928-8

This article was downloaded by **calibre** from https://www.nature.com/articles/d41586-023-01928-8

- NEWS
- 07 June 2023

Morality is declining, right? Scientists say that idea is an illusion

Surveys show people around the world have believed for decades that morals are decaying — but other survey data contradict that perception.

- Mariana Lenharo



Are we not as nice as our forebears? A study questions the idea that morality declines through the generations.Credit: Viacheslav Iakobchuk/Alamy

Adam Mastroianni was always bothered by anecdotal claims that people are becoming less kind, respectful and trustworthy over time. So he took a deep dive into such claims: he wrote a PhD dissertation.

Now Mastroianni and a collaborator have drawn on decades' worth of survey results and other data to find that people around the world have perceived a general moral decline for at least the past 70 years[1]. But the data also show that individuals' evaluation of their contemporaries' morality has remained largely unchanged during that time. Mastroianni, a psychologist at Columbia University in New York City, and his co-author Daniel Gilbert, a psychologist at Harvard University in Cambridge, Massachusetts, conclude that the perception of moral decline is an illusion.

The findings were published in *Nature* on 7 June.

Not so nice

To examine the common idea that morality is waning, Mastroianni and Gilbert analysed US surveys about moral values conducted between 1949 and 2019. In response to around 84% of the survey questions, most participants reported that morality had declined. Similar results were seen in surveys done in 59 other countries.

The authors also conducted their own series of surveys in 2020 and found that US participants said that people were less "kind, honest, nice and good" than they had been in times past, such as the year the participant was born. The perception of moral decline was shared by people of various political ideologies, races, genders, ages and educational levels.



The appetite for right

The researchers also examined surveys that asked people to report on the current state of their own morality and that of their contemporaries. The authors chose surveys that had been administered at least twice, with a minimal interval of ten years, so that answers could be compared across time. These included questions such as “How would you rate the overall state of moral values in this country today?”

If morality had actually declined over time, people would be expected to rate their peers more negatively than had those who took the same survey earlier. But the data revealed that participants’ assessments of their contemporaries’ morality have not changed over time. For Mastroianni, this means that the perception of moral decline is erroneous or that “it’s at least very difficult to find any evidence that this moral decline has happened”.

Melissa Wheeler, an ethics researcher at Swinburne University of Technology in Melbourne, Australia, praised the robustness of the data, noting that it included “responses from millions of archival data points over decades” and a series of well-powered experiments that confirm the findings.

Widespread pessimism

The fact that US residents of all types perceived declining morality surprised psychologist Liane Young, head of the Morality Lab at Boston College in Massachusetts, who was not involved with the study. “That pessimism about human nature is pretty common when people are making judgements about those who are not on their side, on their team. What’s interesting is that this bias seems to be more pervasive and doesn’t seem to be rooted in intergroup bias.”

So why do people think there is this decay? The authors speculate that it has to do with factors such as biased memory; negative memories tend to fade faster than positive ones.

Mastroianni says that the illusion of moral decline might have important societal and even political consequences. For example, a 2015 survey[2] cited in his paper found that 76% of people in the United States agreed that “‘addressing the moral breakdown of the country’ should be a high priority for their government” — a view that could affect voting choices.

“The challenge”, Wheeler points out, “will be in getting people to accept that they hold this illusion, which is so prominent and widely held.”

Nature **618**, 441-442 (2023)

doi: https://doi.org/10.1038/d41586-023-01848-7

References

1. Mastroianni, A. M. & Gilbert, D. T. *Nature* https://doi.org/10.1038/s41586-023-06137-x (2023).

2. Pew Research Center for the People & the Press Research Center Poll: January 2015 Political Survey (Version 2). *Roper Center for Public Opinion Research* https://doi.org/10.25940/ROPER-31096284 (2015).

This article was downloaded by **calibre** from https://www.nature.com/articles/d41586-023-01848-7

- NEWS
- 07 June 2023
- Update 09 June 2023

China’s rolling COVID waves could hit every six months — infecting millions

The latest surge is unlikely to crash the country’s health-care system, but scientists fear hundreds of millions of infections.

- Yvaine Ye



Chinese COVID cases are again climbing as immunity from vaccines and infections wanes.Credit: Daniel Ceng Shou-Yi/ZUMA Press Wire/Shutterstock

The latest surge in COVID-19 cases in China is not surprising to researchers, who say that the country will see an infection cycle every six months now that all COVID-19 restrictions have been removed and highly infectious variants are dominant. But they caution that rolling waves of infection carry the risk of new variants emerging.

“Unfortunately, a new reality with this virus [is that] we will have repeated infections,” says Ali Mokdad, an epidemiologist at the Institute for Health Metrics and Evaluation at the University of Washington in Seattle. “The fear is that this virus will produce a new variant that can compete with the current ones and is more severe.”

The current surge is caused mainly by a highly infectious subvariant of Omicron called XBB.1.5, first identified in India last August. According to Nanshan Zhong, a prominent respiratory physician in China, as many as 65 million people could become infected per week by the of end this month.

This is the first major reinfection wave that China has seen since the central government dropped all its COVID-19 control measures in December, prompting a widespread Omicron outbreak.

China has vaccinated more than 90% of its population, and the outbreak in December infected at least 85% of its people, says Zhong. But immunity is waning, and XBB can evade protection from vaccines and prior infections. Mokdad says that, although XBB has not caused a major rise in hospitalizations and deaths, the sheer number of infections could put pressure on China’s health-care system.

XBB is also causing minor waves in other parts of the world, such as Singapore and the United States.

“This is what we see everywhere, but with a large population such as China, it is more apparent,” says Mokdad.

Yunlong Cao, an immunologist at Peking University in Beijing, and his team have found that antibodies generated against Omicron variants BA.5 and BF.7, the dominant strains during the December wave in China, can provide about four months of protection against strains such as XBB[1].

Kayoko Shioda, an epidemiologist at Boston University in Massachusetts, says that previous COVID-19 surges in other countries have shown that XBB is more transmissible than earlier forms. “Once it enters the population, XBB spreads and becomes the predominant variant much quicker than other variants,” she says.

Last December, more than 200 million people in China contracted COVID-19 in 20 days. This time, the wave is spanning several months, owing to the differences in people’s immune backgrounds, such as antibody levels, says Cao. “The peak of COVID-19 waves will generally become flatter and more stretched out after each cycle, which is a pattern we see in countries like the US. People are still getting infected in the US, just not all at once,” he adds. A flatter wave would also lessen the burden on health-care systems, Cao says.

New boosters

Because China no longer publishes its COVID-19 case count, it is unclear how many people are becoming infected in the latest wave; however, the Beijing health authority says the number of COVID-19 cases reported in the capital city quadrupled between late March and mid-April. Cao says it’s hard to make estimates without accurate data. But on the basis of his past research, he estimates that at least 30% of the population could become reinfected in this wave, amounting to more than 400 million people.

Scientists say having a good surveillance system to monitor and track emerging virus variants is very important, given that infection cycles will continue to happen. A new variant that could supersede current ones remains a concern, says Mokdad. “Imagine a Delta type of variant with XBB capacity of spreading. This will cause us a lot of damage.”

Tracking the virus's evolution also means that scientists can update booster vaccines accordingly.

Facing the ongoing COVID-19 wave, several major Chinese cities, including Beijing and Shanghai, have started to inoculate residents with a quadrivalent COVID-19 booster, made by the Chinese biotechnology company Sinocelltech Group. The vaccine, first approved for emergency use in March, is designed to provide broad-spectrum protection against the Alpha, Beta, Delta and Omicron BA.1 coronavirus variants. Sinocelltech announced last month that one booster shot using this jab can prevent 82% of SARS-CoV-2 infections — including ones caused by XBB — for up to four months. The final results from the vaccine's clinical trial are not yet available.

This week, Chinese state media also announced a new recombinant trivalent vaccine that targets XBB, BA.5 and Delta was approved for emergency use. Describing it as the world's first approved vaccine against XBB, the report said the vaccine will be rolled out within two months.

Although current vaccines can provide good protection against severe disease and death, they are not very good at providing long-lasting protection against infection. "But that doesn't mean we should just give up. Repeated infections, even with a mild virus like XBB, can still lead to health problems like long COVID. Vulnerable people, such as older adults, are still at risk of getting very ill," Cao says.

Nature **618**, 442-443 (2023)

doi: https://doi.org/10.1038/d41586-023-01872-7

Updates & Corrections

- **Update 09 June 2023**: This article has been updated to include a new approval of a vaccine targeting the XBB coronavirus variant.

References

1. Chen, X. *et al. Lancet Reg. Health West Pac.* **33**, 100758 (2023).

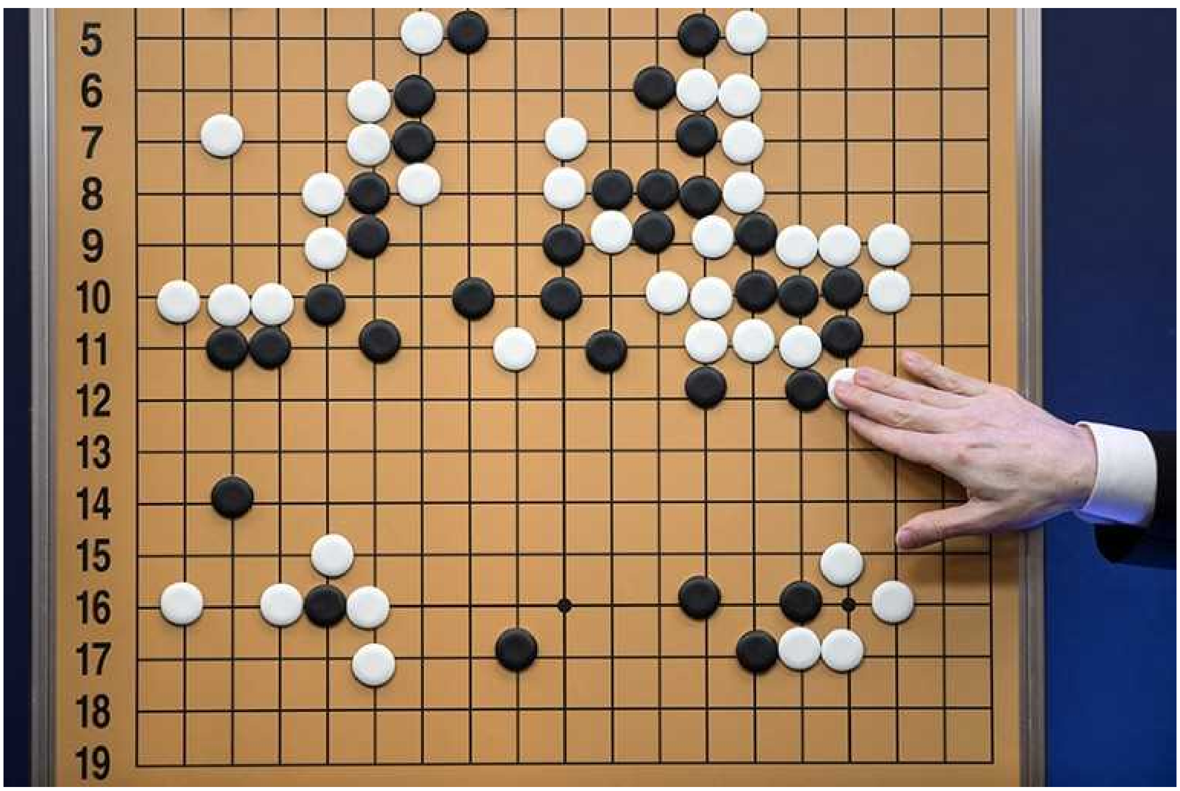
This article was downloaded by **calibre** from https://www.nature.com/articles/d41586-023-01872-7

- NEWS
- 07 June 2023

DeepMind AI creates algorithms that sort data faster than those built by people

The technology developed by DeepMind that plays Go and chess can also help to write code.

- Matthew Hutson



DeepMind used the technology behind AlphaZero — its artificial-intelligence system for playing board games — to generate sorting

algorithms.Credit: Ed Jones/AFP via Getty

An artificial intelligence (AI) system based on Google DeepMind's AlphaZero AI created algorithms that, when translated into the standard programming language C++, can sort data up to three times as fast as human-generated versions.

“We were a bit shocked,” said Daniel Mankowitz, a computer scientist at DeepMind who led the work. “We didn't believe it at first.”



DeepMind AI invents faster algorithms to solve tough maths puzzles

Computer scientists have, for decades, been optimizing how computers sort data to shave off crucial milliseconds in returning search results or alphabetizing contact lists. Now DeepMind, based in London, has vastly improved sorting speeds by applying the technology behind AlphaZero — its artificial-intelligence system for playing the board games chess, Go and shogi — to a game of building sorting algorithms. “This is an exciting result,” said Emma Brunskill, a computer scientist at Stanford University, California.

The system, AlphaDev, is described in a paper in *Nature*[1], and has invented faster algorithms that are already part of two standard C++ coding libraries, so are being used trillions of times per day by programmers around the world.

Starting small

The researchers first applied AlphaDev to the task of sorting numbers by size. They started small, with algorithms that sorted only 3, 4, or 5 numbers at a time, but these are important because they're used by algorithms that sort longer lists. AlphaDev operated at the level of assembly instructions: code generated by automated compilers from code that programmers write in C++, before it is translated into the 1s and 0s of machine code.

AlphaDev works similarly to its predecessor, AlphaZero, which combines computer versions of deliberation and intuition to choose moves in a board game[2]. AlphaDev doesn't choose moves; instead, it chooses instructions to add to a procedure (in what DeepMind engineers call AssemblyGame).

When using deliberation, at each decision point, AlphaZero considers its possible moves, its possible moves after each of those moves, and so on, in a branching fashion, calculating which moves are most likely to end with a win. But considering all possible branches might take longer than the age of the universe, so it uses something akin to intuition to narrow its search. At each step, the computer program feeds the game state into neural networks — complex, tunable mathematical functions — that highlight the most promising moves. During training, it continually updates the networks based on game outcomes. It also explores moves by not always picking the one that is currently highest-rated.

Rewards offered

AlphaDev can take one of four types of actions that involve comparing values, moving values between locations, or jumping to a different part of the program. After each step, it tries sorting a set of lists and receives a reward for how many items in the lists it sorted correctly. It plays until it sorts all the lists perfectly or reaches a program length limit, then starts a new program from scratch.

The neural networks evaluated and rewarded the programs not only on correctness, but also on speed. Mankowitz's team trained the system to evaluate speed on the basis of either the total number of instructions or the processing time. Depending on the processor used and the number of values to be sorted, AlphaDev's best algorithms took between 4% and 71% less

time than did human algorithms. But when the algorithms were called multiple times to sort lists of one-quarter of a million values, the cumulative time saving was only 1–2%, because of other code it did not optimize.



[AI maths whiz creates tough new problems for humans to solve]

The DeepMind team also applied AlphaDev to non-sorting algorithms. Its version of an algorithm used to convert data stored in a particular format into bytes took 67% less time than a standard version. And its hashing algorithm, used in data storage and retrieval, took 30% less time than a standard one.

To see where AlphaDev eked out its gains, the team took a closer look at its algorithms. For sorting, they found two new tactics, which they called the AlphaDev swap move and the AlphaDev copy move. Mankowitz compares them to ‘Move 37’, a surprising move that AlphaGo, a predecessor to AlphaZero, made against the human Go champion Lee Sedol in 2016, at an exhibition match in Seoul. “It’s something that, in hindsight, was actually fundamental to winning the game and influenced how we thought about strategies,” he says.

In terms of the science, “I don’t know that there’s anything particularly deep in there,” says Michael Littman, a computer scientist at Brown University in Providence, Rhode Island, who notes that AlphaZero has already been around for six years. “But the engineering is phenomenal.” He adds that the researchers behind DeepMind are good at fitting the method to new

problems. Last year, DeepMind also modified AlphaZero to create AlphaTensor[3], which invented faster ways to multiply grids of numbers.

In future, the DeepMind team would like to apply AlphaZero-style algorithms to more kinds of problems, even the design of hardware itself, Mankowitz says. “We really want to be tackling the whole stack.”

Nature **618**, 443-444 (2023)

doi: https://doi.org/10.1038/d41586-023-01883-4

References

1. Mankowitz, D. L. *et al. Nature* **618**, 257–263 (2023).
2. Silver, D. *et al. Science* **362**, 1140–1144 (2018).
3. Fawzi, A. *et al. Nature* **610**, 47–53 (2022).

This article was downloaded by **calibre** from https://www.nature.com/articles/d41586-023-01883-4

- NEWS
- 02 June 2023

Colombia’s ‘cocaine hippo’ population is even bigger than scientists thought

The most comprehensive census yet reveals that there could be twice as many of the invasive animals than previous estimates indicated.

- Luke Taylor



Wild descendants of the hippo population introduced in Colombia by Pablo Escobar.Credit: Raul Arboleda/AFP via Getty

Colombia's invasive hippo population is even larger than researchers had thought, according to the most thorough census of the animals conducted yet. Scientists were already concerned about the hippos — considered the largest invasive animal in the world — threatening native plants and animals in the country, and had been calling for drastic measures to reduce the population. The census results have only heightened that fear.



Pablo Escobar's 'cocaine hippos' spark conservation row

A few years ago, researchers estimated how fast the animals were reproducing, to project that about 98 hippos were living along the country's Magdalena River and its tributaries in 2020[1]. But the new study, for which a research team counted the animals in person, by drone and using other tracking methods, estimates that there are 181–215 of them residing in Colombia.

"Before, one argument against dealing with the hippos was that our information was limited and our arguments theoretical," says ecologist Rafael Moreno, who participated in the study while at the Alexander von Humboldt Biological Resources Research Institute in Bogotá. "But we have put that argument to bed now. This study shows that this is a real issue, and that the state must act urgently."

By drone and on foot

Colombia's 'cocaine hippos' are all descendants of three females and one male illegally imported by drug-cartel leader Pablo Escobar. After he died in 1993, the hippos (*Hippopotamus amphibius*) escaped from his estate and established themselves in the Magdalena River. Without the natural predators or droughts of their native Africa to keep them in check, the giant herbivores have bred rapidly to form the largest population of the animals outside that continent.

Colombian officials have struggled to manage the hippos. After they ordered an aggressive male to be killed in 2009, a photo of soldiers posing with the corpse sparked outrage and halted efforts to rein them in. Some communities now depend on the tourism the hippos bring, whereas others — particularly fishing communities — live in fear of the highly territorial animals, which can weigh 3 tonnes and can tear off a person's limbs or trample them.



Landmark Colombian bird study repeated to right colonial-era wrongs

Colombia's environment ministry commissioned the census to get a better picture of the problem and how to manage it. Completing it was challenging: despite the hippos' large size, it is difficult to find and count them

accurately. They are nocturnal, immerse themselves in water for 16 hours a day and roam large distances.

Comprising researchers at the National University of Colombia in Bogotá, the Humboldt Institute and Cornare, an environmental body managing an area where the hippos live, the team made various trips in 2021 and 2022 by car, boat and on foot to count the hippos. In locations where the animals could not be reached safely, they used drones to count them or footprints to estimate the population.



One way of managing the hippo population is by contraception, administered using a dart.Credit: Cornare via Reuters

The researchers found that 37% are juveniles, indicating that the animals are breeding rapidly. One hypothesis for this is that the hippos are reaching sexual maturity earlier than they do in Africa owing to the lush conditions in Colombia. Another is that the animals are having greater reproductive success because there are fewer fights among them for territory and resources, Moreno says. But evidence is needed to confirm the actual cause.

While counting the hippos, the team documented the myriad ways in which the animals are damaging Colombia's ecosystems. As they waddle their huge bodies to and from rivers, the hippos are eroding riverbanks and carving out muddy paths that divide forests. Moreover, they are outcompeting other animals for habitat and resources, with the West Indian manatee (*Trichechus manatus*), neotropical otter (*Lontra longicaudis*) and capybara (*Hydrochoerus hydrochaeris*) most under threat.

Exploring solutions

With serious attacks on humans in 2020 and 2021, and a car crash leaving a hippo dead on the highway in April, solutions are needed, scientists say.

One strategy currently being tried is to administer contraceptives to the animals by dart. That might eventually get rid of the hippos by preventing their reproduction, but it is slow, costly and hasn't been tested on hippos on such a scale before. A modelling study published in April[2] estimated that this method could eradicate the hippos in 45 years at a cost of at least US$850,000.



The world's species are playing musical chairs: how will it end?

Another strategy — capturing, anaesthetizing and transporting the hippos by helicopter to a facility to be castrated — would cost at least $530,000 and take up to 52 years for eradication, the study found. Both calculations are underestimates, given that the models were fed earlier, lower estimates of the hippo population, before the census results had been released.

Meanwhile, the regional agency responsible for dealing with the hippos is strapped for funding and relying on donations of contraceptives from the United States. Agency officials are negotiating with countries, including India and Mexico, to export some of the animals to sanctuaries abroad. But this would cost $3.5 million, according to those working on the strategy.

There is no single best solution, says Olga Montenegro, a biologist at the National University who worked on the census. The hippos live in various groups, some of which might be easily relocated, but others are well established and are breeding quickly.

Many researchers advocate culling the animals. They say it is the swiftest, most humane thing to do and that it would solve the problem before it becomes impossible to fix. The cost of killing the hippos must be weighed against that of losing native flora and fauna in Colombia — the second-most biodiverse country in the world — they add.

Nature asked the environment ministry how it would manage the growing hippo population, given the new findings, but it did not comment.

“There is a moral weight to the decision to cull a hippo. But the weight of the other decision — inaction — is far greater,” Moreno says. “I hope this is something the politicians will understand.”

Nature **618**, 444-445 (2023)

doi: https://doi.org/10.1038/d41586-023-01818-z

References

1. Castelblanco-Martínez, D. N. *et al. Biol. Conserv.* **253**, 108923 (2021).

2. Subalusky, A. L. *et al. Sci. Rep.* **13**, 6193 (2023).

This article was downloaded by **calibre** from https://www.nature.com/articles/d41586-023-01818-z

| Section menu | Main menu |

- NEWS
- 06 June 2023

‘Science was heard’: woman who was convicted of killing her children pardoned after inquiry

Kathleen Folbigg, who was jailed in Australia in 2003 over the sudden deaths of her four young children, has been pardoned and released in the wake of new scientific evidence.

- Dyani Lewis



Kathleen Folbigg appears during an inquiry in 2019.Credit: Joel Carrett/AAP Image via Alamy

Researchers and legal experts are applauding the role of an independent scientific adviser in an Australian judicial inquiry whose evidence led to the pardon and release of a woman sent to prison 20 years ago. Kathleen Folbigg, jailed in 2003 over the deaths of her four young children, walked free on 5 June after the inquiry concluded there was "reasonable doubt as to the guilt of Ms Folbigg for each of the offences for which she was originally tried". Scientists involved in the inquiry are now calling for legal reform so that there is a formal process in Australia to present emerging scientific evidence.

"Science was heard in this case," says Michael Toft Overgaard, a protein scientist at Aalborg University in Denmark who was an expert witness at the inquiry. "The whole case has been a bit surreal for us as scientists to be part of."

Genetic evidence not available at the time of Folbigg's conviction revealed that rare mutations in the protein calmodulin could have caused the deaths of her two daughters, and a neurogenetic disorder could have led to the death of one of her sons. Calmodulin controls calcium concentration in cells, and among other roles helps to regulate the heart's contractions.

Regarding her pardon, "I'm so happy for her," says geneticist Carola Vinuesa at the Francis Crick Institute in London. Vinuesa was approached by Folbigg's lawyers in 2018 to sequence Folbigg's genome and the genomes of her deceased children to see whether there was a genetic explanation for their sudden deaths, which occurred between 1989 and 1999.

Scientific adviser

Vinuesa identified mutations in a gene called calmodulin 2 in Folbigg's genome as well as in those of her two daughters, which could have explained the girls' deaths. But that evidence was not enough to convince the commissioner of the first inquiry, who upheld Folbigg's convictions. So in 2019, Vinuesa contacted the Australian Academy of Science, which

gathered support to petition the governor of New South Wales to grant Folbigg a pardon, based on subsequent work by Overgaard and others that showed how the mutations impaired protein function.

At that stage, every other legal avenue had been exhausted. The governor ordered a new inquiry, and its commissioner, former state chief justice Thomas Bathurst, appointed the academy to act as a scientific adviser.

In that role, the academy recommended scientific expert witnesses and advised on the scope of expertise for each witness. The academy put forward some 30 researchers, around half of whom presented evidence at the inquiry. Other experts were also called by the prosecution and defence teams. The academy's chief executive, Anna-Maria Arabia, says the inquiry heard "the most up-to-date science from the most qualified scientists, wherever they were in the world". The expert witnesses were independent of both the prosecution and defence, says Arabia, and available for all parties to interrogate.

Peter Schwartz, a cardiologist at the Italian Auxological Institute in Milan, Italy, and a world leader in calmodulin mutations that cause sudden death, was one of the expert witnesses. He has advised on nearly 40 medico-legal cases, mostly in the United States, and says that having independent advice from the academy helped to ensure relevant experts presented world-leading evidence to the inquiry, instead of relying on one or two local experts. "I don't recall ever having seen that in a trial of this kind," he says, and "it goes to the credit of the Australian justice system."

Overgaard says that he and other experts were given time to provide the necessary background so that lawyers thoroughly understood the science. On one occasion, he spent more than five hours explaining how mutations in the calmodulin protein could impair its function. The inquiry was also put on hold at one stage so that Overgaard and his team could update their evidence with results of further experiments they ran to address another expert's questions, he says.

Time for reform

Jason Chin, a legal academic at the Australian National University in Canberra, says that an independent scientific adviser could be used more often, especially in judicial inquiries that have greater discretion on process.

Arabia says that the case demonstrates how the science and justice systems can work together, and that it should prompt law reforms to create a more “science-sensitive legal system”. She and others are calling for the establishment of a criminal case review commission, similar to that in the United Kingdom, which can revisit cases when there are advances in the science and new evidence comes to light.

“I do hope they succeed,” says Overgaard, “so that experts with specific knowledge relevant to the case are called.”

But despite praise for the process, researchers caution that science doesn’t necessarily make a case black and white. “When the science is really nuanced, and really new and evolving as this was, you still may not get consensus,” says Hugh Watkins, a cardiologist at the University of Oxford, UK, who studies genes that cause sudden cardiac death.

Watkins adds: “It really does come down to the legal assessment of where that residual uncertainty leaves you.”

Nature **618**, 445-446 (2023)

doi: https://doi.org/10.1038/d41586-023-01871-8

This article was downloaded by **calibre** from https://www.nature.com/articles/d41586-023-01871-8

- NEWS
- 06 June 2023

Tanzania’s researchers offered US$22,000 to publish in international journals

The move is designed to encourage research and help boost institutions’ rankings. But some researchers say it will reward those already established in their careers.

- Syriacus Buguzi



Tanzania wants to improve the global visibility of its scientists and the position of universities in international rankings.Credit: Getty

Tanzania's government has offered scientists 50 million Tanzanian shillings (roughly US$22,000) if they can publish their research in a well-known journal. Researchers had until the end of last month to apply for the scheme.

A total of 1 billion shillings ($423,575) has been allocated annually for the Research Excellence Award. It aims to boost publications from Tanzania's researchers in "internationally renowned and the most reputable journals", says Maulilio Kipanyula, director of science, technology and innovation at Tanzania's Ministry of Education, Science and Technology in Dar es Salaam.

Researchers could apply for one of the awards for publications in the natural sciences, mathematics or medicine in the financial year 2022–23. Kipanyula told *Nature* that Tanzania is following South Africa, Ireland, Australia and Pakistan in offering cash for publications. Three years ago, China ended the practice after many decades of paying researchers for publishing in well-known journals.

Paula Stephan, an economist at Georgia State University in Atlanta who studies how monetary incentives affect publishing, says "it does not come as a surprise to see that Tanzania has joined the list of countries offering such incentives to faculty".

To be eligible, researchers will have had to publish papers in the highest 10% of journals, as measured by impact factor in selected disciplines. Papers could be original research; secondary data analyses or systematic reviews; or meta-analyses. "This will improve global rankings of Tanzanian scientists and higher-learning institutions," Kipanyula adds.

More than 80% of Tanzania's research publications are the result of international collaborations, according to a 2019 report commissioned by the UK government. The award guidelines state that, in such cases, only one eligible author will be allowed to submit an application. If there is more than one Tanzanian author, they will be encouraged to share the prize.

Some have welcomed the scheme as a much-needed incentive for scientists to conduct research in a country with limited resources, where most academics are focused on teaching and where the average monthly salary for

a professor is around 1,990,000 shillings, about $844. But others say the funds could have been better spent.

Tanzania spent roughly 0.5% of its gross domestic product on research and development in 2022. Moreover, the country relies on foreign aid for more than 50% of its national research expenditure.

“Getting a publication accepted in high-impact-factor journals requires a lot of effort,” says Frank Kagoro, a Tanzanian epidemiologist and African Regional Manager at the Global Health Network at the University of Oxford, UK. “Initiatives like this may encourage scientific development, rigour and competition which can also be a catalyst to address local issues and improve science and development,” he explains.

However, Haruna Kanaabi, a human-rights researcher at Makerere University in Kampala, Uganda, argues that awards for papers in well-known journals are most likely to reward people who are already established in their careers, whereas Tanzania’s government should be encouraging those at earlier career stages. Celebrating those “that have made it up at the top is a nice gesture but misplaced”, he says.

Aneth David, a microbiologist at the University of Dar es Salaam, agrees. She adds that the government needs to ease institutional bureaucracy and allow Tanzania’s researchers more freedom to choose their areas of study.

“It is no secret that laws and regulations around research and scholarly communication have become tighter over the past few years and have stifled academic freedom,” she says.

“There won’t be high-quality papers in high-impact journals if scientists are not free to choose what they study and be free to do so based on scientific principles, not politicians’ whims,” she adds. “A sustainable internal funding mechanism would go a long way to allow researchers to pursue their curiosity, generate knowledge and eventually publish.”

Stephan says that the initiative seems to differ from those of other countries in that it requires the publications to address national research areas stipulated by public-sector agencies such as the Tanzanian Commission for

Science and Technology. This, she says, raises questions such as “to what extent will high-impact journals be interested in publishing research in [Tanzania’s] national research areas”.

Nature **618**, 446 (2023)

doi: https://doi.org/10.1038/d41586-023-01729-z

This article was downloaded by **calibre** from https://www.nature.com/articles/d41586-023-01729-z

- NEWS FEATURE
- 14 June 2023

Lab mice go wild: making experiments more natural in order to decode the brain

Armed with technologies to track a creature's every move, neuroscientists are gaining insights into animal — and human — behaviour.

- Kerri Smith



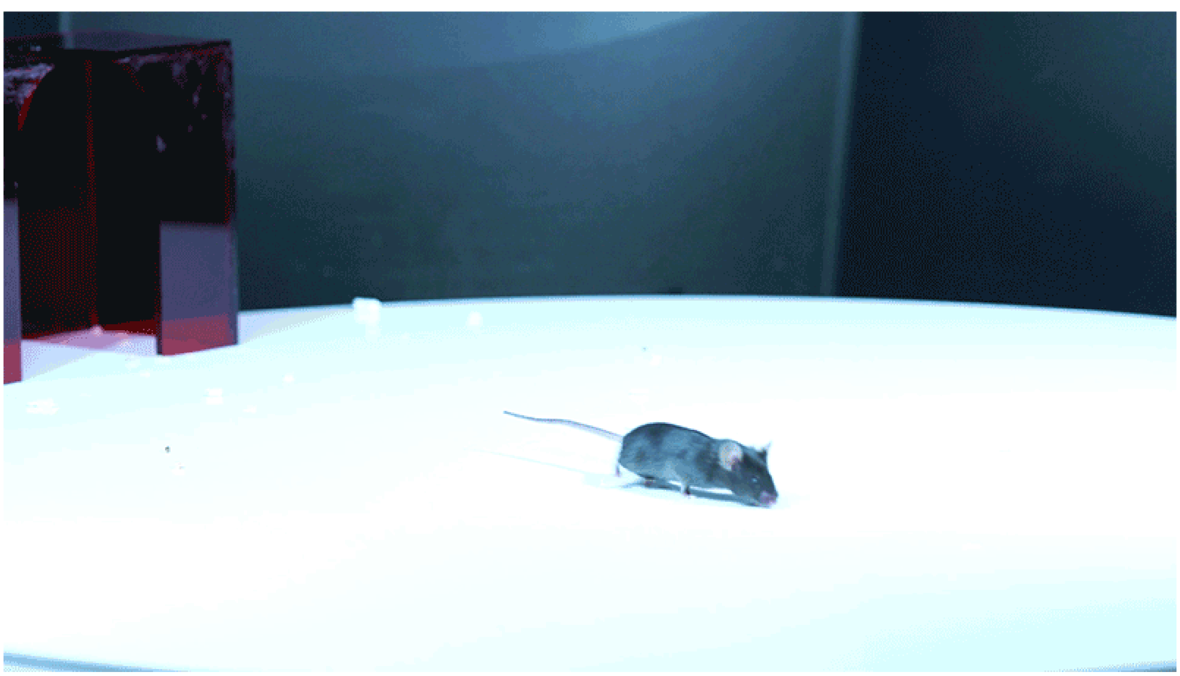
Illustration by Piotr Kowalczyk

In a dimly lit laboratory in London, a brown mouse explores a circular tabletop, sniffing as it ambles about. Suddenly, silently, a shadow appears.

In a split second, the mouse's brain whirs with activity. Neurons in its midbrain start to fire, sensing the threat of a potential predator, and a cascade of activity in an adjacent region orders its body to choose a response — freeze to the spot in the hope of going undetected, or run for shelter, in this case a red acetate box stationed nearby.

From the mouse's perspective, this is life or death. But the shadow wasn't cast by a predator. Instead, it is the work of neuroscientists in Tiago Branco's lab, who have rigged up a plastic disc on a lever to provoke, and thereby study, the mouse's escape behaviour. This is a rapid decision-making process that draws on sensory information, previous experience and instinct.

Branco, a neuroscientist at the Sainsbury Wellcome Centre at University College London, has wondered about installing a taxidermied owl on a zip wire to create a more realistic experience. And his colleagues have more ideas — cutting the disc into a wingspan shape, for instance. "Having drones — that would also be very nice," says Dario Campagner, a researcher in Branco's lab.



Scientists evoke the escape response in a mouse. (The ‘predator’ shadow has been enhanced for visibility).Credit: Dario Campagner and Tiago Branco

The set-up is part of a growing movement to step away from some of the lab experiments that neuroscientists have used for decades to understand the brain and behaviour. Such exercises — training an animal to use a lever or joystick to get a reward, for example, or watching it swim through a water maze — have established important principles of brain activity and organization. But they take days to months of training an animal to complete specific, idiosyncratic tasks. The end result, Branco says, is like studying a “professional athlete”; the brain might work differently in the messy, unpredictable real world.

Mice didn’t evolve to operate a joystick. Meanwhile, many behaviours that come naturally — such as escaping a predator, or finding scarce food or a receptive mate — are extremely important for the animal, says Ann Kennedy, a theoretical neuroscientist at Northwestern University in Chicago, Illinois. They are “critical to survival, and under selective pressure”, she says. By studying these natural actions, scientists are hoping to glean lessons about the brain and behaviour that are more holistic and more relevant to everyday activity than ever before.



Habit-linked brain circuits light up in people with eating disorders

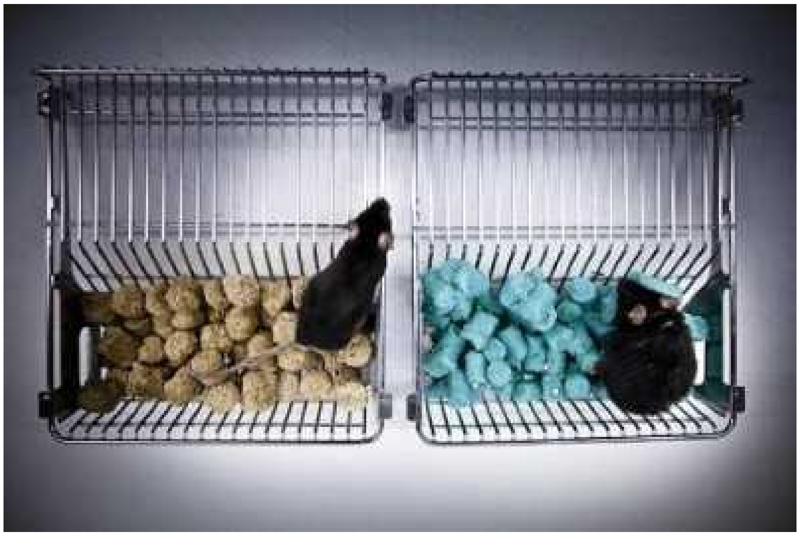
As neuroscientists continue to hone their naturalistic set-ups using the latest technologies for brain imaging and behaviour tracking, they are finding better, more nuanced ways to use animals to study pain responses and conditions such as Down's syndrome and autism. Others are rethinking popular theories about aggression and fear. And some are looking for ways in which these methods could enable richer studies of human behaviour, says Sarah Lisanby, a psychiatrist who directs the Division of Translational Research at the US National Institute for Mental Health (NIMH) in Bethesda, Maryland. That could be a game-changer for research into some psychiatric conditions.

"Until we understand the brain basis of what goes on for individuals when they are experiencing symptoms, we'll continue to not effectively serve them," says Lisanby, who in the past year helped to launch more than US$25 million in funding for research to quantify natural behaviours in humans and other animals.

Researchers in the field acknowledge that there is a lot to learn about their new set-ups, and the utility of the approach is an open question. "Do we really learn more by letting animals do what they want?" asks Sandeep Robert Datta, a neuroscientist at Harvard Medical School in Boston, Massachusetts, who studies naturalistic mouse behaviour. "Collectively we're just getting started."

Back to the wild

The approach takes cues from early ethologists, who studied natural behaviour through detailed diary entries and long-running activity logs of their favourite birds or insects. Dutch biologist Niko Tinbergen, considered one of the founders of ethology, would spend hours sitting on sand dunes on the North Sea coast jotting down descriptions of gull behaviour. (He won the 1973 Nobel Prize in Physiology or Medicine for his contributions to ethology, alongside two other scientists.)



Sex, food or water? How mice decide

These biologists studied natural behaviour in some depth, but “they never got to the brain, because they couldn’t”, says Branco. By contrast, scientists who have been able to explore the brain have a limited palette of behaviours that are accessible for study.

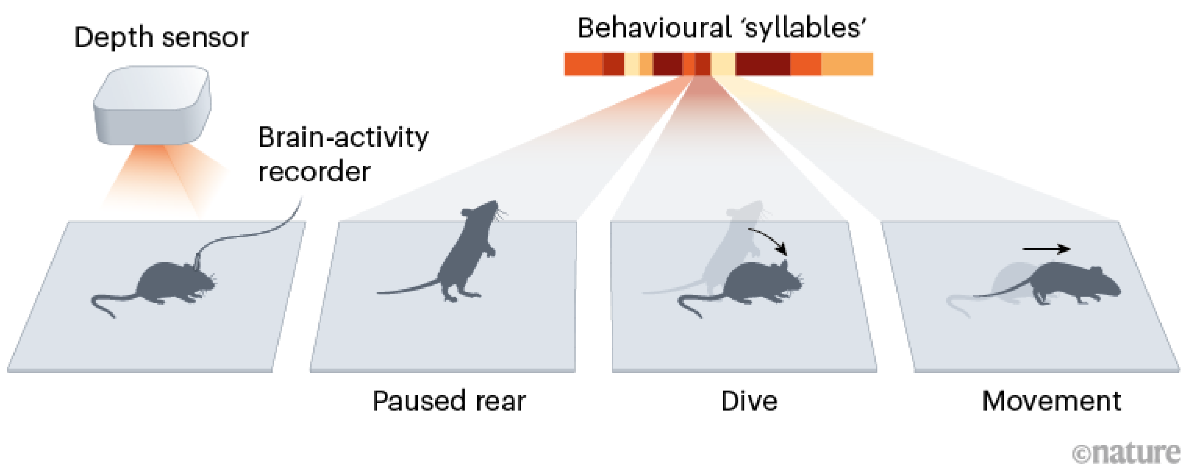
Some teams, such as Branco’s, are studying specific behaviours while recording brain activity using electrodes mounted on animals’ heads. Downstairs from his team’s fake-predator experiments, for instance, researchers at the Sainsbury Wellcome Centre are tracking animals as they seek food. The team has built an arena with little wheels embedded in the floor. When a mouse digs at the wheels, cereal pellets pop out at a rate chosen by the scientists. The goal is to reproduce the variability of food sources in the wild.

Other labs let the mice mind their own business, and catalogue what they do from moment to moment. Even seemingly simple behaviour has a complexity that is interesting to neuroscientists, says Datta. “Running in an empty bucket in the dark represents a significant cognitive challenge to the animal,” he says. Using 3D imaging, Datta and his colleagues have catalogued a ‘grammar’ of mouse body language, broken into simple modular actions or ‘syllables’ such as rearing up on the hind legs or bobbing the head[1]. They have used it to look at how different stimuli or genetic

manipulations might change the patterns of actions (see ‘Cataloguing behaviours’).

CATALOGUING BEHAVIOURS

Using cameras that track movement in three dimensions, scientists have catalogued dozens of spontaneous mouse behaviours, or ‘syllables’, and the sequences in which they typically occur. Behaviours such as rearing, diving and moving around can then be matched to brain activity in some experimental set-ups.



Source: Ref. 1

The computing power and analytical techniques needed to record, automatically track and catalogue hours of behaviour have become available only in the past five years — aided by machine learning. This has set the field in motion. For example, neuroscientist [Mackenzie Mathis at the Swiss Federal Institute of Technology in Lausanne] debuted an open-source, motion-tracking software package called DeepLabCut[2] in 2018, which has been installed 500,000 times. And other, similar packages have emerged.

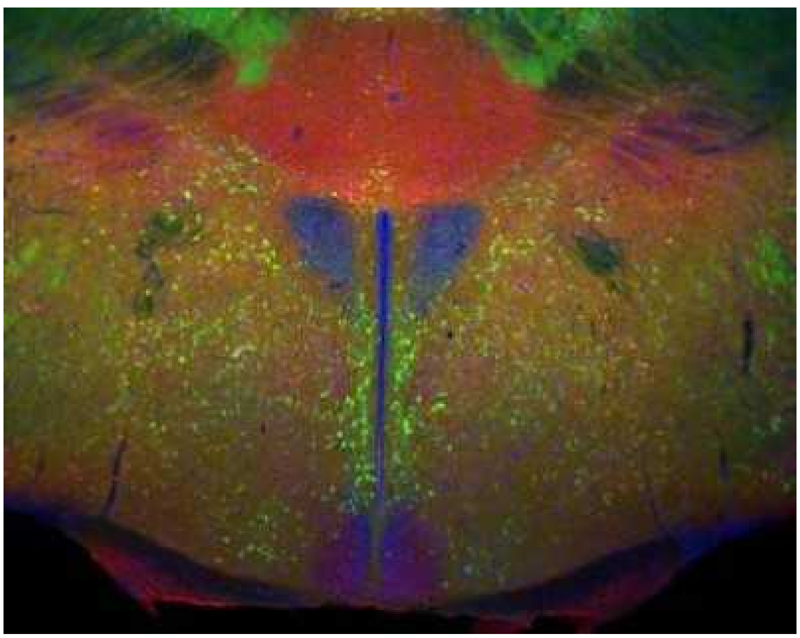
“We have the tools to answer these big questions about what gives rise to behaviour, which is so fundamental to who we are,” says Mathis. “I don’t see a limit.”

Already, the young field is producing a slew of insights into the diversity of animal behaviour and how the brain creates it.

Act natural

Some discoveries question long-held ideas about why animals behave the way they do. For instance, researchers have often suggested that the behaviour of female animals is made more variable by hormones that govern the menstrual cycle, such as oestrogen. But when Datta's team left mice to explore an environment, the team found that, overall, female behaviour was less variable than that of males, and that oestrogen levels had little effect on behaviour[3].

Studying mice fighting each other led Kennedy and her colleagues — who at the time were members of neurobiologist David Anderson's lab at the California Institute of Technology in Pasadena — to rehabilitate a classic ethological theory that had fallen out of favour. The idea, from Austrian ethologist Konrad Lorenz, who shared the 1973 Nobel prize with Tinbergen, was that an internal signal of aggression builds up until it tips the animal into action. Some critics thought the model was overly simplistic, lacked a neurobiological basis and didn't account for learning.



'Mirror neurons' fire up during mouse battles

Kennedy and her colleagues, however, found evidence for the process in mice[4]. When a mouse interacted with another, a group of neurons in the hypothalamus gradually ramped up activity to a level that correlated with what the mouse did. At low levels, the animal might freeze or ignore the

other mouse. But as levels built up over tens of seconds, it might show signs of aggression, such as trying to mount the other mouse. At high levels, mice began to attack others outright. The neurons seem to function like a volume dial for aggression — and what's more, the levels vary between mice, with some creatures almost never attacking and some quick to do so[4].

Other experiments refine previous theories of how the brain controls the body. Decades of studies had suggested that a region called the amygdala governs fear, "and that everything that deals with defensive response has to involve the amygdala", says Branco. When his team looked at mice escaping from a predator, they found a circuit that acts as a shortcut from a mouse's eyes to the back of its brain, where it can initiate the movement for escape[5]. Later on, the mouse's amygdala might help it to learn from the experience — but the mouse doesn't need this region in the moment.

Pause over pain

Some research into natural behaviours might already have clinical implications. Ishmail Abdus-Saboor at Columbia University in New York City and his colleagues use naturally behaving mice to study pain, hoping to build a better picture of its causes and potential treatments than standard models allow. "If you see the doctor for pain, it's not because you sat at home and someone poked you in a constrained area," he says. It's because walking hurts or lying in bed gives you back pain. "We haven't been measuring that in animal models."

In 2021, a team led by Abdus-Saboor and Victoria Abraira, a neuroscientist at Rutgers University in New Brunswick, New Jersey, published a preprint showing that a commonly used anti-inflammatory painkiller called meloxicam seemed to work fine when tested in a standard pain assay, but not when the mice were behaving spontaneously. This suggested that the animals were still in pain but that conventional tests were missing something[6].



[Chronic pain can be treated — so why are millions still suffering?]

In both set-ups, mice were given an injection in their paw that causes inflammation, and then received the painkiller. In the standard test for sensitivity to pain, the scientists subjected the paw to heat and saw that the mice had a minimal response, suggesting the drug was working.

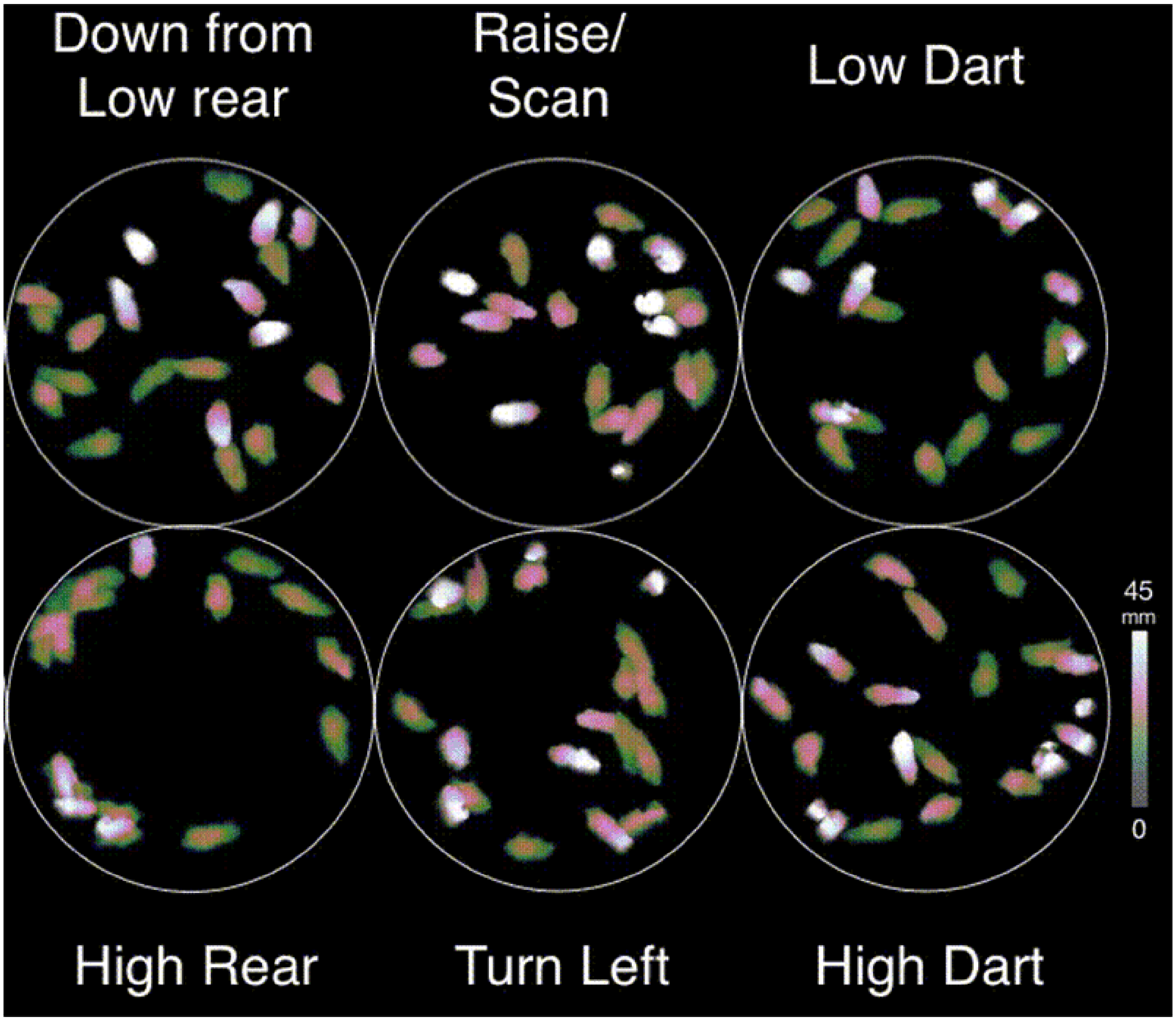
But when the team observed mice behaving spontaneously, they noticed that certain actions, such as rearing onto the hind legs, persisted — the mice were still showing symptoms of pain. “This was quite provocative and a little shocking and alarming,” says Abdus-Saboor. Spontaneous behaviours revealed a more complex side of the pain response, such as the way an injured foot can change the way a person walks. Perhaps animal models are not capturing the full experience of pain, says Abdus-Saboor — which could explain why painkillers that seem to work in rodents have often failed in human trials.

Abdus-Saboor consults for the US pharmaceutical companies Eli Lilly and Doloromics, which are both considering adopting naturalistic assays in drug discovery. Other groups are also working on naturalistic models of pain.

Better science

Studying natural behaviour has produced some interesting findings — but it could also improve neuroscience more broadly.

At the NIMH, neuroscientist Yogita Chudasama directs the Rodent Behavioural Core, a centralized facility that helps researchers across the National Institutes of Health (NIH) to characterize rat and mouse behaviour. Her unit is setting up equipment that will allow researchers to collect data on spontaneous behaviours over long time periods to decrease variability in experiments and make the findings more reliable. In a typical experiment, an animal might be lifted from its home cage and taken somewhere else to be observed. But this new environment could affect how it behaves. Longer-term observation flattens out variables that might be affecting the animal. Furthermore, an environment with less human interference will allow the rodent to behave more naturally than it would in a constrained setting. The next step would be to integrate brain-activity recording.



Researchers use 3D recordings and machine learning to parse different types of spontaneous mouse behaviour.Credit: Datta Lab, Department of Neurobiology, Harvard Medical School

Some researchers using the core facility are monitoring animals for the long term, to compare those in which a particular gene has either been left intact or mutated in a way that disrupts its function. Some of the changes that result can be subtle. “We believe that by looking over long timescales, we will see nuances of behavioural impairments,” says Chudasama.

Vivek Kumar, a neurogeneticist at The Jackson Laboratory in Bar Harbor, Maine, is also looking for ways to improve animal models. He has been studying behaviour in animal models of Down’s syndrome and autism — conditions that come with cognitive changes that are hard to reproduce in animals. But by looking instead at movement, which is much simpler to do,

Kumar has found that animals with gene variants related to these conditions display differences in gait[7]. If these differences are caused by the same genes or circuits as the cognitive changes, says Kumar, then an intervention for one could influence the other. The team hopes to use tests of motor behaviours to screen hundreds of compounds for their effects on gait.

Cash injection

The interest in natural behaviour is spreading from animals to humans, and with it lots of cash. The NIH will dish out $20 million in 2024 and 2025, as part of its BRAIN Initiative, for researchers who wish to develop systems that track behaviour and brain activity in humans. “There is a lack of access to what the brain is doing when people engage in complex behaviour,” says Lisanby, who helped to develop this and other funding opportunities to support naturalistic neuroscience.

Human psychiatric conditions such as obsessive compulsive disorder, can manifest in the lab and be studied in a scanner. But most episodes occur at home, where it is a huge challenge to monitor brain activity while people are mobile. Lisanby hopes that the NIH fund will help researchers to develop tools for measuring the brain and behaviour outside the lab. This might include sensors that people can wear at home, and mobile brain-recording devices that are better than those currently available.

In such a new field, there are plenty of teething problems. Abdus-Saboor says that it can be hard to find researchers who have the mix of skills required. “Most experimentalists are not trained in mathematics, computation, computer science and coding. So we do find a disconnect,” he says. Last year, he and others started a course for graduate students and postdoctoral researchers at The Jackson Lab on quantifying behaviour.

Many scientists have ambitious plans for the young field. A lot of teams, including Branco’s, have dreams of tracking more than one animal, and over timescales of weeks to months. They hope to understand how the brain chooses between behaviours, to monitor the social dynamics of groups and even to study how brains lay down memories or plan for the future. To track multiple animals, neural recording would need to be wireless to prevent

cables from getting tangled and use advanced algorithms to track movements. Existing systems can have trouble distinguishing between animals when they interact and overlap, especially if they are similar sizes and colours. Mathis wants to catalogue mouse behaviour over an animal's lifetime, and use the information to create 'digital twin' mouse models to be used as a reference.

Researchers acknowledge that conventional approaches are not going away. Although they are excited by the new wave of technologies, they are realistic about how much the latest methods can achieve. "These tools aren't magic. They're just a pair of glasses," says Datta. "They help us look."

Nature **618**, 448-450 (2023)

doi: https://doi.org/10.1038/d41586-023-01926-w

References

1. Wiltschko, A. B. *et al. Neuron* **88**, 1121–1135 (2015).
2. Mathis, A. *et al. Nature Neurosci.* **21**, 1281–1289 (2018).
3. Levy, D. R. *et al. Curr. Biol.* **33**, 1358–1364 (2023).
4. Remedios, R. *et al. Nature* **550**, 388–392 (2017).
5. Evans, D. A. *et al. Nature* **558**, 590–594 (2018).
6. Bohic, M. *et al.* Preprint at bioRxiv https://doi.org/10.1101/2021.06.16.448689 (2021).
7. Sheppard, K. *et al. Cell Rep.* **38**, 110231 (2022).

This article was downloaded by **calibre** from https://www.nature.com/articles/d41586-023-01926-w

- NEWS FEATURE
- 02 June 2023
- Correction 02 June 2023

‘It’s a vote for hope’: first gene therapy for muscular dystrophy nears approval, but will it work?

The FDA’s decision, expected this month, follows several setbacks and delays and will pose difficult choices for the families of children with Duchenne muscular dystrophy.

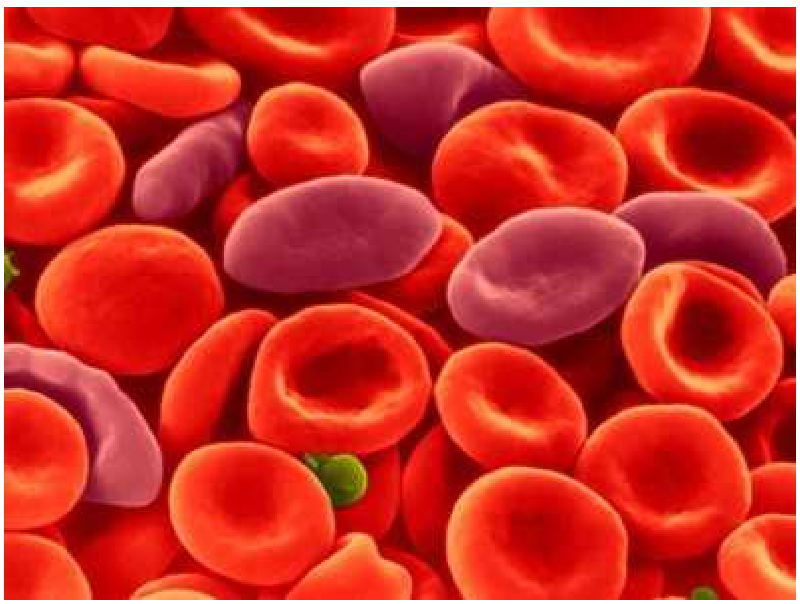
- Sara Reardon



The parents of children with Duchenne muscular dystrophy have been pushing for swift approval of therapies. Credit: Eric Kruszewski/*New York Times*/Redux/eyevine

The road to gene therapies for genetic disorders has been long — and expensive — but the field could soon get some good news. On 22 June, the US Food and Drug Administration will decide whether to grant a fast-track approval to the first gene therapy for Duchenne muscular dystrophy (DMD), a genetic disorder that affects around 1 in 3,500 boys. Children with DMD can't make a protein called dystrophin, resulting in progressive muscle degeneration and death in their twenties due to heart or respiratory failure.

The therapy, known as SRP-9001, is made by Sarepta Therapeutics based in Cambridge, Massachusetts. It would be the 13th gene therapy that the FDA has approved since 2017, and the first to target a prevalent genetic disease in children. The accelerated approval would allow the drug to reach the market before large clinical trials have been completed, on the basis of evidence that the therapy allows boys to make an engineered form of dystrophin.



Gene therapy's comeback: how scientists are trying to make it safer

The decision date was delayed late in May after FDA officials and advisers raised concerns about the strength of Sarepta's data so far; SRP-9001 seems to have only a modest effect on muscle function, and only in some people. A

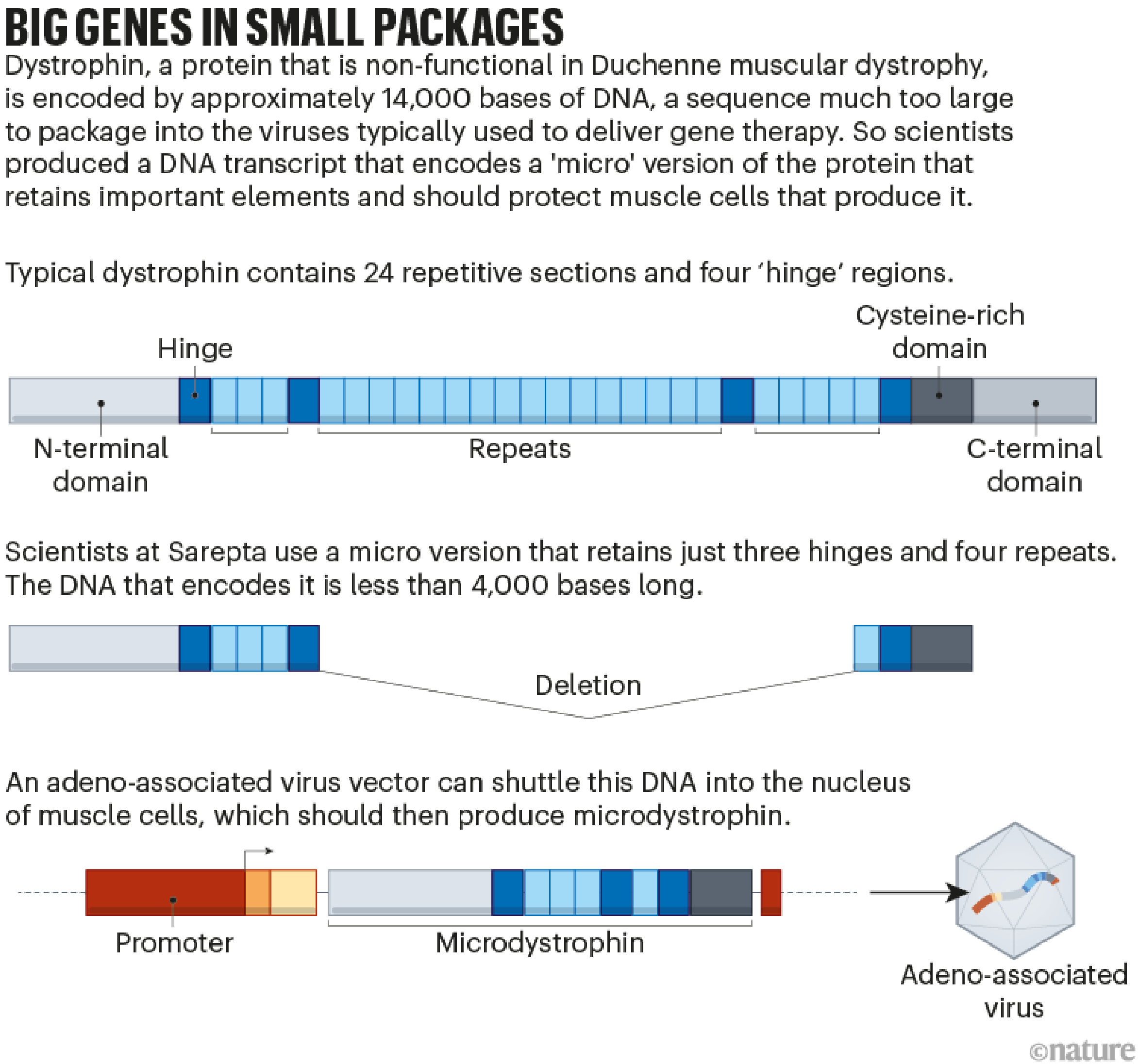
press release from the company says that the agency is likely to approve it only for boys aged four and five, but might expand that range if an ongoing clinical trial shows that the drug's efficacy is worth the risks associated with it. Analysts predict that the one-time treatment will cost US$2 million. A Sarepta spokesperson declined to reveal the price tag until the drug is approved but says that it will be priced "below the value it will bring to patients".

Still, some scientists hope that the approval will pave the way for more gene therapies. "To have a big success like this is going to validate the technology quite a bit," says Jeffrey Chamberlain, a neurologist at the University of Washington in Seattle who developed some of the technologies used by Sarepta. "It's a way to get our foot in the door now and help kids."

Why has DMD been so challenging to treat?

DMD seems like a straightforward target for gene therapy. The disease affects boys almost exclusively, because they have only one copy of the X chromosome, where the dystrophin gene is located; girls with a disease-causing variant have a backup copy. Replacing even some working protein in muscle cells should reverse the disease, or at least halt its progression.

But developing that replacement has proved difficult. Dystrophin is the longest gene in the human genome and is much too large to fit into the adeno-associated virus (AAV) vector commonly used to deliver gene therapies. Sarepta and several other companies have got around this by designing a gene that encodes just the most important parts of the protein (see 'Big genes in small packages'). The resulting 'microdystrophin' is only partly effective.



Getting the gene into enough cells to make a difference has also been challenging. Muscle makes up 30–40% of the body's mass, so treating DMD requires extremely high doses of the AAV vector. This raises the risk of side effects, including organ damage. And because muscle cells divide frequently — especially in a growing child — the amount of the gene in the body will be diluted over time. It's still unclear how long the therapy's effect will last, although Sarepta's data from four children suggest that the clinical improvements can persist four years after treatment.

How effective is the treatment?

Although most FDA approvals require drugs to pass two phase III, placebo-controlled clinical trials testing the efficacy of the treatment, the accelerated approval pathway that Sarepta is seeking allows companies to rely on 'surrogate endpoints' to prove their value. These measures show a biological effect — in this case, high levels of microdystrophin — rather than an improvement in symptoms.

Sarepta's first large phase III trial is still under way, with data expected to be released late this year. The FDA has instead evaluated several midstage trials in a total of 52 boys aged four to seven. These showed that the boys' muscles made microdystrophin, but the children had no statistically significant improvement in muscle function one year after the drug was administered.



Will the FDA change how it vets drugs following the Alzheimer's debacle?

Reporting by the news organization *STAT* and others found that FDA staff planned to reject Sarepta's application. But Peter Marks, director of the FDA's Center for Biologics Evaluation and Research (CBER) intervened — scheduling a public meeting with an independent scientific advisory committee on 12 May, to weigh in on whether the drug should receive an accelerated approval. This committee narrowly recommended the approval by eight votes to six. Although the FDA doesn't have to follow the advice of the committee, it typically does.

The advisory committee evaluated evidence for surrogate markers as well for clinical benefits. Some of those who voted in favour of the therapy say that because DMD progresses slowly, it might be difficult to determine the drug's effect after just one year. But when Sarepta looked only at the 32 boys younger than six years old who received SRP-9001 or a placebo, the treated boys' muscle function was significantly improved.

Sarepta's chief scientific officer, Louise Rodino-Klapac, doesn't think that the drug actually functions better in these younger boys. Rather, she says, the older boys in the placebo control group were healthier at the beginning of the study than their younger counterparts, making the drug's effects look weaker in this group by comparison. Rodino-Klapac expects that the ongoing phase III trial, which better controlled for each patient's initial condition, will show that it works equally well in older boys' muscles.

The statistically significant improvement in the younger boys was enough to convince Donald Kohn, an FDA advisory committee member who does stem-cell research at the University of California Los Angeles, to vote in favour of the drug. "From a statistical standpoint, [Sarepta] haven't pled their case," he says. But, he adds, the efficacy in younger boys suggests that the drug is having a real effect. "It's a vote for hope," says Kohn.

Others are more concerned by what looks like an overall lack of efficacy. "I understand the appeal of 'Something is better than nothing', but the question is at what cost," says Caleb Alexander, an epidemiologist at Johns Hopkins University in Baltimore who was also a committee member and voted against the drug. Not only can SRP-9001 have significant side effects, he says, but the placebo-controlled trials completed so far do not strongly suggest that high levels of microdystrophin even predict clinical outcomes. "No number of post-hoc analyses of subgroups of boys change that fact," he says.

The FDA and advisory committee members also raised concerns about whether Sarepta would finish its clinical trial, which is still evaluating more than 60 boys who had initially received the placebo and are now receiving the drug. They pointed out that the company did not complete trials of three previous DMD drugs that the FDA had approved on an accelerated basis.

They are still on the market. A spokesperson for Sarepta says that those trials have now enrolled enough patients and are now under way.



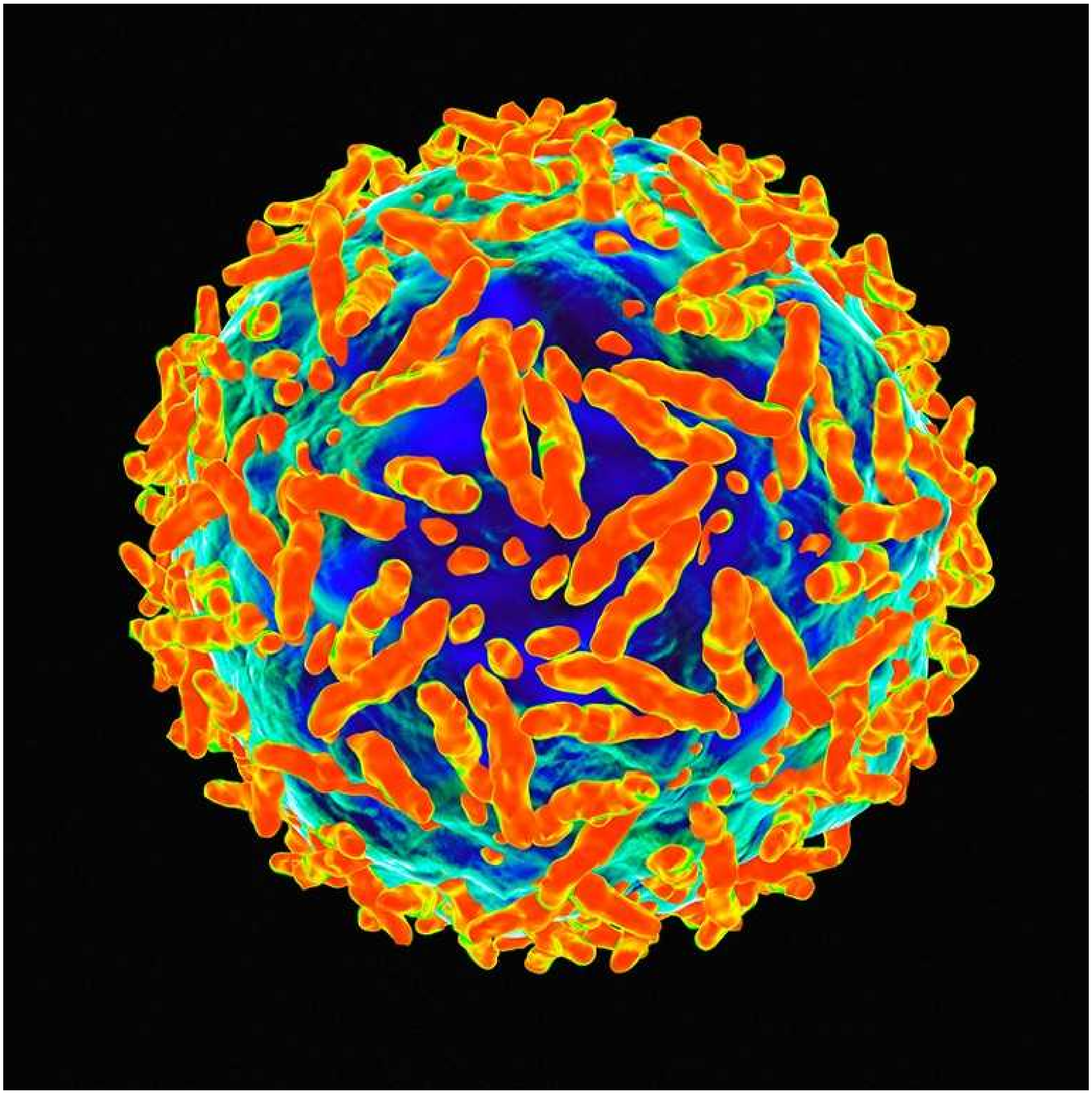
[Expensive treatments for genetic disorders are arriving. But who should foot the bill?]

“It’s a really complex decision. I don’t envy the FDA,” says Sharon Hesterlee, chief research officer of the Muscular Dystrophy Association in Chicago, Illinois. Although she agrees that the therapy’s effect might be modest and carries significant risks, “the risk of doing nothing is 100% fatal”.

What are the risks of DMD gene therapy?

As well as the risks from the large amounts of AAV required to get the microdystrophin gene into muscles, the gene itself seems to cause severe side effects in some children. The FDA had suspended multiple DMD gene-therapy trials owing to contamination concerns and patients becoming seriously ill. And in 2021, the agency temporarily halted a study by Pfizer, based in New York City, which is testing its own version of microdystrophin, after a patient died.

To investigate the problem, Pfizer, Sarepta and Solid Biosciences, based in Charlestown, Massachusetts, decided to pool their microdystrophin data. Their study, which will soon be published in the *New England Journal of Medicine*, found that certain mutations in the dystrophin gene cause the immune system to recognize microdystrophin as a foreign invader and attack it, causing dangerous inflammation in the muscle and heart. The companies began genetically screening clinical-trial participants and excluding those with the relevant mutations — less than 5% in Sarepta's case — which eliminates the problem for now. But Carsten Bönnemann, a neurologist at the National Institute of Neurological Disorders and Stroke in Bethesda, Maryland, who co-led the study, says that gene therapies will eventually have to be tailored to help these patients.



Gene therapies delivered by adeno-associated viruses can only be administered once in a patient.Credit: Mehau Kulyk/Science Photo Library

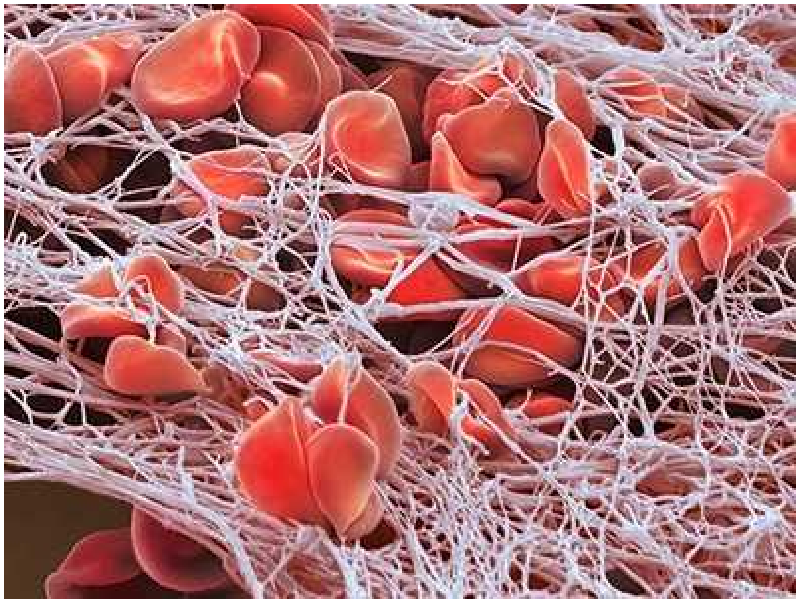
Still, Hesterlee notes that even those who are eligible for Sarepta's newly approved therapy have a tough choice ahead of them. AAV gene therapies can be administered only one time in a person's life: once the immune system has encountered a virus vector, it is likely to attack it in the future. And because AAV is the vector of choice for most gene therapies, parents of a child with DMD might have to choose whether to treat their child now, with the only approved therapy, or wait in the hope that Sarepta or another company will release something better in the future. That means allowing

the disease to progress, and potentially stabilizing it only after boys have lost more muscle function. “A few months can be a lot for families,” Hesterlee says.

What other DMD gene therapies are on the horizon?

Pfizer is conducting a phase III trial in 99 boys with their version of microdystrophin in a similar AAV vector and expects to release its initial results next year. Several other companies are developing their own microdystrophins, including Regenxbio in Rockville, Maryland, which launched a small trial earlier this year.

Others are searching for ways to improve the viral vector that carries the gene. In 2022, Solid Biosciences halted its phase II trial of its own microdystrophin formulation to switch to a different vector that specifically targets muscle cells and can probably be given in lower doses. The company plans to begin dosing patients later this year. And at the American Society for Gene & Cell Therapy’s annual meeting in Los Angeles in May, Chamberlain presented a new system involving several AAV vectors that each carry a portion of the dystrophin gene. The gene products can then link together to create a version of dystrophin that is much larger than a single vector could carry. Further on the horizon are approaches that will infuse patients with genes encoding the CRISPR–Cas9 genome editing system with the intention of permanently altering the dystrophin gene itself in muscle cells.



[Researchers welcome $3.5-million haemophilia gene therapy — but questions remain]

And other researchers are looking into alternative vectors, including lipid capsules, nanoparticles, antibodies that carry dystrophin genes to muscle cells and viruses that integrate into a patient's genome. These treatments, at least in theory, could be administered to people who have already received AAV gene therapy.

But with an approved therapy now available, companies trying to test their own therapies could find it more difficult to recruit trial participants, who might not want to risk getting a placebo. Furthermore, health-services researcher Reshma Ramachandran at Yale University in New Haven, Connecticut, worries that if companies know that they can get a microdystrophin therapy approved, they will be more likely to create their own slightly tweaked versions instead of spending a longer time developing innovative approaches of their own.

How will this approval affect other gene therapies?

Gene-therapy development has always been a risky and expensive business, and the recent economic downturn has seen many biotechs shut down.

Nevertheless, the number of gene-therapy approvals by the FDA has risen each year since its first one — Kymriah for leukemia — in 2017.

Approval for a DMD gene therapy would still represent a milestone, however. Until now, most approved gene therapies have addressed cancers, extremely rare diseases and conditions such as retinal disorders, which are easy to target with a virus. DMD is different, Chamberlain says, both in the number of technical challenges that researchers had to address and in the large number of patients that the therapy could serve. The agency is now poised to consider more than a dozen gene and cell therapies this year, including two for sickle-cell disease — a genetic condition that is much more prevalent than DMD.

The FDA seems to be prioritizing these approaches. In recent talks, CBER director Marks said that the agency plans to lean more heavily on accelerated approval pathways for gene therapies, including broader use of surrogate endpoints. “We can’t be so careful about our approvals under accelerated approval that we prevent potentially life-saving therapies from getting to market in a timely manner,” he said in a January speech at the Muscular Dystrophy Association’s annual meeting.

That concerns some scientists, who worry that patients will end up paying high prices for ineffective drugs. “I do think the notion of regulatory flexibility can get bent out of shape,” Alexander says. Ramachandran points to the FDA’s 2021 accelerated approval of an Alzheimer’s drug, aducanumab, that successfully eliminated amyloid plaques in patients’ brains but did not measurably improve their health. The FDA subsequently approved another, similar drug that showed only minor clinical benefit. “Patients with Duchenne’s deserve not just more options, they need better options,” Ramachandran says.

But Chamberlain contends that even a small amount of success will help the whole field. “There’s tremendous room for improvement,” he says. “I think this is going to buy us some time to make those improvements.”

Nature **618**, 451-453 (2023)

doi: https://doi.org/10.1038/d41586-023-01799-z

Updates & Corrections

- **Correction 02 June 2023**: The US Food and Drug Administration told Sarepta that it would complete its review by 22 June, not 27 June as previously reported.

This article was downloaded by **calibre** from https://www.nature.com/articles/d41586-023-01799-z

Books & Arts

- **Particle, wave, both or neither? The experiment that challenges all we know about reality** [13 Jun 2023]
  • Thomas Young's double-slit experiment originally served to prove that light is a wave — but later quantum versions have made for a much fuzzier picture.
- **Why incentives work — or don't, and more: Books in brief** [14 Apr 2023]
  • Andrew Robinson reviews five of the best science picks.

- ESSAY
- 13 June 2023

Particle, wave, both or neither? The experiment that challenges all we know about reality

Thomas Young's double-slit experiment originally served to prove that light is a wave — but later quantum versions have made for a much fuzzier picture.

- Anil Ananthaswamy



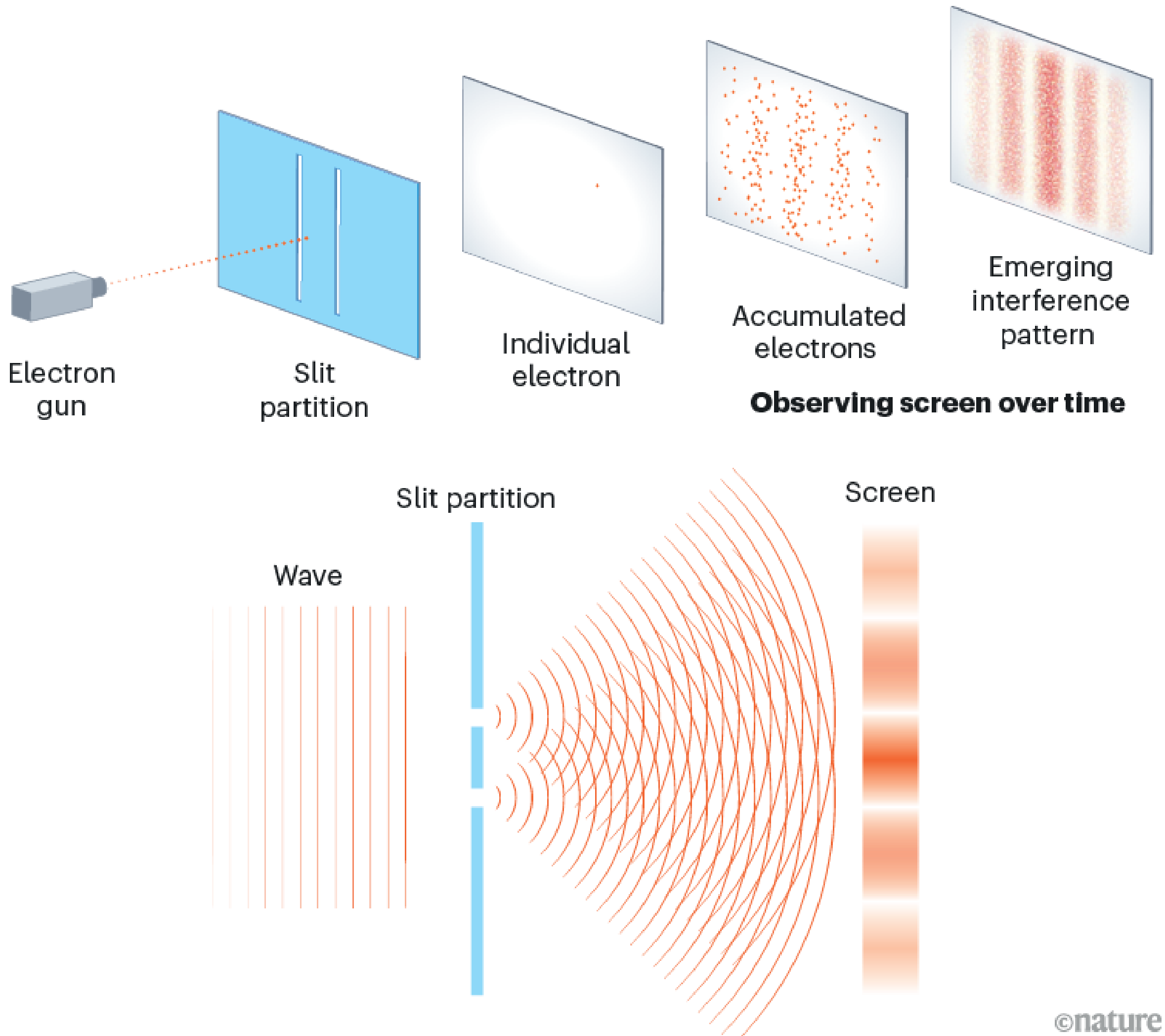
The double-slit experiment's interference patterns suggest something is in two places at once.Credit: Huw Jones/Getty

Thomas Young, born 250 years ago this week, was a polymath who made seminal contributions in fields from physics to Egyptology. But perhaps his most enduring legacy is proving Isaac Newton wrong about light — and igniting a debate about the nature of reality that still persists.

"The experiments I am about to relate", he told the Royal Society of London[1] on 24 November 1803, "may be repeated with great ease, whenever the sun shines." In a simple, modern form, Young's 'double-slit' experiment involves shining light of a single frequency (say, from a red laser) through two fine, parallel openings in an opaque sheet, onto a screen beyond. If light were made of streams of particles, as Newton conjectured, you would expect to see two distinct strips of light on the screen, where the particles pile up after travelling through one slit or the other. But that's not what happens. Instead, you see many bands of light and dark, strung out in stripes like a barcode: an interference pattern (see 'Wave–particle weirdness').

WAVE–PARTICLE WEIRDNESS

When quantum objects such as electrons are fired one by one through a pair of closely spaced slits, they behave like particles: each one hits a screen placed on the far side at exactly one point. But they also behave like waves: successive hits build up a banded interference pattern exactly like that generated by a wave passing through the slits (bottom). This wave–particle duality is described by a mathematical tool known as the wavefunction.



Interference is possible only if light behaves as a wave that strikes both slits at once and diffracts through each, creating two sets of waves on the other side of the slits that propagate towards the screen. Where the crest of one wave overlaps with the crest of the other, you get constructive interference and a patch of light. Where a crest meets a trough, you get destructive interference and darkness.

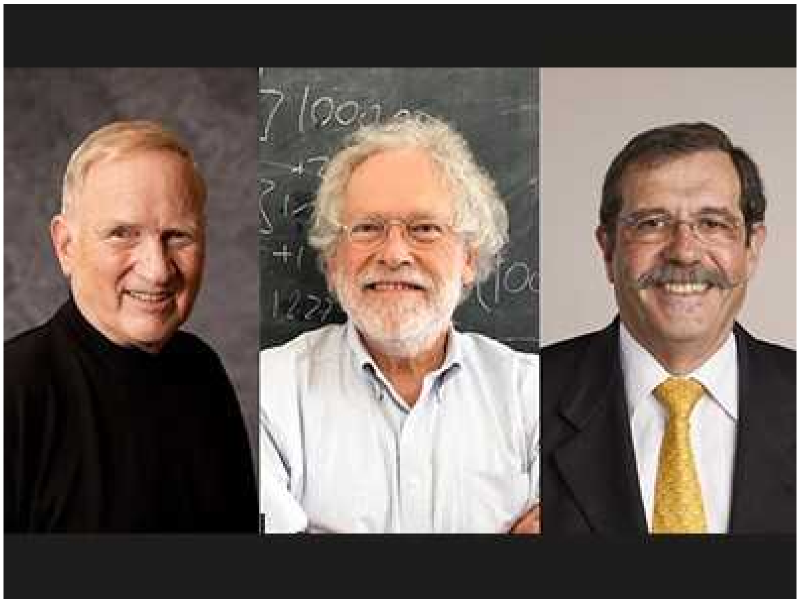
It’s hard to overstate how wild this discovery was to physicists in Young’s time. But the wildness truly began when Max Planck and Albert Einstein laid the foundations for quantum mechanics in the early twentieth century. Today, quantum mechanics forms a peerlessly accurate framework to explain the basic elements of material reality and their interactions. Pretty early on, it became clear that it implied that light is made of indivisible units of energy called photons — particles, in fact. The amount of energy each carried was proportional to the frequency of the light. Some carry enough of a wallop to knock electrons off atoms of metal, giving us the photoelectric effect that enables today’s solar cells. (It was the study of this effect that led Einstein to his conclusions about light’s particulate nature.)

With the emergence of quantum mechanics, the idea of light as a wave faced a challenge. But it wasn’t as simple as going back to the particle view. Further tests of quantum theory using the double-slit experiment only deepened the mystery. And it hasn’t been solved yet.

Singularly quantum

Imagine, now, that your light source can shoot individual photons of red light at the two slits, while guaranteeing that only one photon goes through the apparatus at any time. A photographic plate on the other side records where the photons land. Classical intuition says each photon can go through only one slit or the other. So, this time, we should see photons accumulating over time and forming two strips of light on the photographic plate. Yet the mathematics of quantum theory implied that the interference pattern would persist.

It was several decades before the technology matured enough to verify these predictions experimentally, using more complex set-ups that were in principle the double-slit. At first, it wasn’t done with photons, but with electrons — entities that we know as particles, but that quantum mechanics predicts act as waves, too. Then, in the 1980s, a team led by Alain Aspect at the Optical Institute in Palaiseau, France, performed the double-slit experiment with single photons[2]. Quantum theory won out: an interference pattern emerged, even when only single particles passed through the slits.



'Spooky' quantum-entanglement experiments win physics Nobel

Aspect won a share of the 2022 Nobel prize in physics for his contribution to confirming the predictions of quantum mechanics through experiment. But such experiments leave matters of interpretation wide open. There is simply no way to comprehend what's happening with minds attuned to the classical world of everyday objects.

When it comes to the double-slit experiment, quantum mechanics does tell a form of story. It says that a photon's position is described by a mathematical abstraction called the wavefunction — which, as the name suggests, behaves like a wave. This wavefunction, mathematically speaking, hits the two slits, diffracts into two sets of waves and recombines to create the interference pattern. The value of the wavefunction at any location on the photographic plate lets you calculate the probability of finding the photon there. The probability is very high in regions of constructive interference, and very low in regions of destructive interference.

In a sense, then, a photon or any other quantum object acts like both a particle and a wave. This 'wave–particle duality' embodies many of the central conceptual mysteries of quantum mechanics that are unresolved to this day. Even if you could know everything about a photon's initial state, there's no way to tell exactly where it'll land on the detector. You have to talk in terms of probabilities given by the wavefunction. These probabilities

are borne out only when thousands or tens of thousands of photons are sent through the double slit, one by one.

Before the measurement — in this case, detection by the photographic plate — the mathematics says the particle exists in a superposition of states: in a sense, it has taken both paths, through the right slit and the left. Standard quantum mechanics says that the wavefunction 'collapses' when measured, and that the act of observation in some way precipitates that collapse. Before this, the photon has a finite probability of being found in many different regions, but on measurement, the wavefunction peaks at the location in which the photon appears (the probability there equals 1) and is nullified everywhere else (probability equals 0).

It gets even odder. If you can determine which path the photon took on its way to the detector, it acts like a particle that does indeed go through one slit or the other: the interference pattern disappears. But if you cannot glean this 'which-way' information, the photon acts like a wave. Whenever there are two or more ways for a photon — or, indeed, any quantum object — to get to a final state, quantum interference occurs.

What's a wavefunction?

But to generate interference, something has to go through — or at least interact in some way with — both slits. In the mathematics, the wavefunction does the job. Some physicists would say that the wavefunction simply represents information about the quantum system and is not real — in which case it's hard to explain what interacts with both slits at once. But you can explain the interference pattern if you consider the wavefunction to be real.

This creates its own problems. Imagine a real wavefunction that spreads for kilometres and kilometres before an observer detects the photon. At this point, the wavefunction peaks at the photon's location, and simultaneously drops to zero everywhere else — over a large, macroscopic distance. This suggests a kind of instantaneous, non-local influence that bothered Einstein no end. One can avoid this with interpretations of quantum theory that don't collapse the wavefunction, but that opens other cans of worms.



Thomas Young reported the results of the original double-slit experiment in 1803.Credit: Photo Researchers/SPL

Perhaps the most notorious is the many-worlds interpretation, the brainchild of US physicist Hugh Everett in the 1950s. This argues that every possible event — in the case of the double slit, a particle going through the left and the right slit — happens, each in its own world. There is no collapse: measurement simply reveals the state of the quantum system in that world.

Detractors ask how it's possible to justify this constant proliferation of worlds, and how, in a many-worlds framework, you can explain why measuring quantum systems yields probabilities, given that there are always definite outcomes in each world.

The de Broglie–Bohm theory, named after quantum pioneers Louis de Broglie and David Bohm, provides another alternative. It says that particles exist with definite positions and momenta, but are guided by an all-encompassing, invisible 'pilot' wave, and it's this wave that goes through both slits. The most profound implication of this theory, that everything is linked to everything else in the Universe by the underlying pilot wave, is one many physicists have trouble accepting.

In the 1970s and 1980s, physicists upgraded the double-slit experiment to seek clarity about the nature of quantum reality, and the perplexing role observation apparently has in collapsing a defined, classical reality out of it. Most notably, John Wheeler at the University of Texas at Austin designed the 'delayed choice' thought experiment[3]. Imagine a double-slit set-up that gives the option of gathering or ignoring information about which way the particle went. If you ignore the 'which-way' information, you get wave-like behaviour; if you don't, you get particle-like patterns.



Superconducting qubits cover new distances

With the apparatus on the 'collect which-way information' setting, send a photon through the double slits. It should act like a particle and go through one slit or the other. But just before the photon lands on the detector, flip the apparatus to ignore the which-way information. Will the photon, until then supposedly a particle, suddenly switch to being a wave?

Decades later, Aspect's team performed this experiment with single photons and showed that the answer is yes[4]. Even if the photon had ostensibly travelled through the entire set-up as a particle, switching the apparatus setting so that it ignored which-way information caused it to act like a wave. Did the photon travel back in time and come back through the two slits as a wave? To avoid such nonsensical explanations, Wheeler argued that the only way to make sense of the experiment was to say that the photon has no reality — it's neither wave nor particle — until it's detected.

Back in the 1980s, Marlan Scully, then at the University of New Mexico in Albuquerque, and his colleagues came up with a similarly befuddling thought experiment[5]. They imagined collecting the which-way information about a photon by using a second photon 'entangled' with the first — a situation in which measuring the quantum state of one tells you about the quantum state of the other. As long as the which-way information can in principle be extracted, the first photon should act like a particle. But if you erase the information in the entangled partner, the mathematics showed, the first photon goes back to behaving like a wave. In 2000, Scully, Yoon-Ho Kim and their colleagues reported performing this experiment[6]. Surprisingly — or unsurprisingly, by this stage — intuition was once again defeated and quantum weirdness reigned supreme.

Larger and still larger

Others are still pushing the double slit in new directions. This year, Romain Tirole at Imperial College London and his colleagues described an experiment in which the slits were temporal: one slit was open at one point in time and the second slit an instant later[7]. A beam of light that goes through these temporal slits produces an interference pattern in its frequency spectrum. Again, the mathematics predicts exactly this behaviour, so physicists aren't surprised. But it is more proof that the double-slit

experiment highlights the lacunae in our understanding of reality, a quarter of a millennium after the birth of the man who devised it.

The double-slit experiment's place in the pantheon of physics experiments is assured. But it would be further cemented if and when physicists using it were able to work out which theory of the quantum world is correct.

For example, some theories posit that quantum systems that grow bigger than a certain, as-yet-undetermined size randomly collapse into classical systems, with no observer needed. This would explain why macroscopic objects around us don't obviously work according to quantum rules — but how big does something have to be before it stops acting in a quantum way?

In 2019, Markus Arndt and Yaakov Fein at the University of Vienna and their colleagues reported sending macromolecules called oligoporphyrins, composed of up to 2,000 atoms, through a double slit to see whether they produce an interference pattern[8]. They do, and these patterns can be explained only as a quantum phenomenon. Arndt's team and others continue to push such experiments to determine whether a line exists between the quantum and the classical world.

Last year, Siddhant Das at the Ludwig Maximilian University of Munich, Germany, and his colleagues analysed the double-slit experiment in the context of the de Broglie–Bohm theory[9]. Unlike standard quantum mechanics, this predicts not just the distribution of particles on the screen that leads to the spatial interference pattern, but also the distribution of when the particles arrive at the screen. The researchers found that their calculations on the distribution of arrival times agreed qualitatively with observations made two decades before, in a double-slit experiment using helium atoms[10]. But it was difficult to prove their case definitively. They are awaiting better data from a similar double-slit experiment done with current technology, to see whether it matches predictions.

And so it goes on, a world away from anything Young or his peers at the Royal Society could have conceived of more than two centuries ago. "Thomas Young would probably scratch his head if he could see the status of today's experiments," says Arndt. But that's because his experiment, so simple in concept, has left us scratching our heads to this day.

Nature **618**, 454-456 (2023)

doi: https://doi.org/10.1038/d41586-023-01938-6

References

1. Young, T. *Phil. Trans. R. Soc.* **94**, 1–16 (1804).
2. Grangier, P., Roger, G. & Aspect, A. *Europhys. Lett.* **1**, 173 (1986).
3. Miller, W. A. & Wheeler, J. A. *Foundations of Quantum Mechanics in the Light of New Technology* (Eds Nakajima, S., Murayama, Y. & Tonomura, A.) 72–84 (World Scientific, 1997).
4. Jacques, V. *et al. Science* **315**, 966–968 (2007).
5. Scully, M. O. & Drühl, K. *Phys. Rev. A.* **25**, 2208 (1982).
6. Kim, Y.-H. *et al. Phys. Rev. Lett.* **84**, 1 (2000).
7. Tirole, R. *et al. Nature Phys*. https://doi.org/10.1038/s41567-023-01993-w (2023).
8. Fein, Y. Y. *et al. Nature Phys.* **15**, 1242–1245 (2019).
9. Das, S., Deckert, D.-A., Kellers, L. & Struyve, W. Preprint at https://arxiv.org/abs/2211.13362 (2022).
10. Kurtsiefer, Ch., Pfau T. & Mlynek, J. *Nature* **386**, 150–153 (1997).

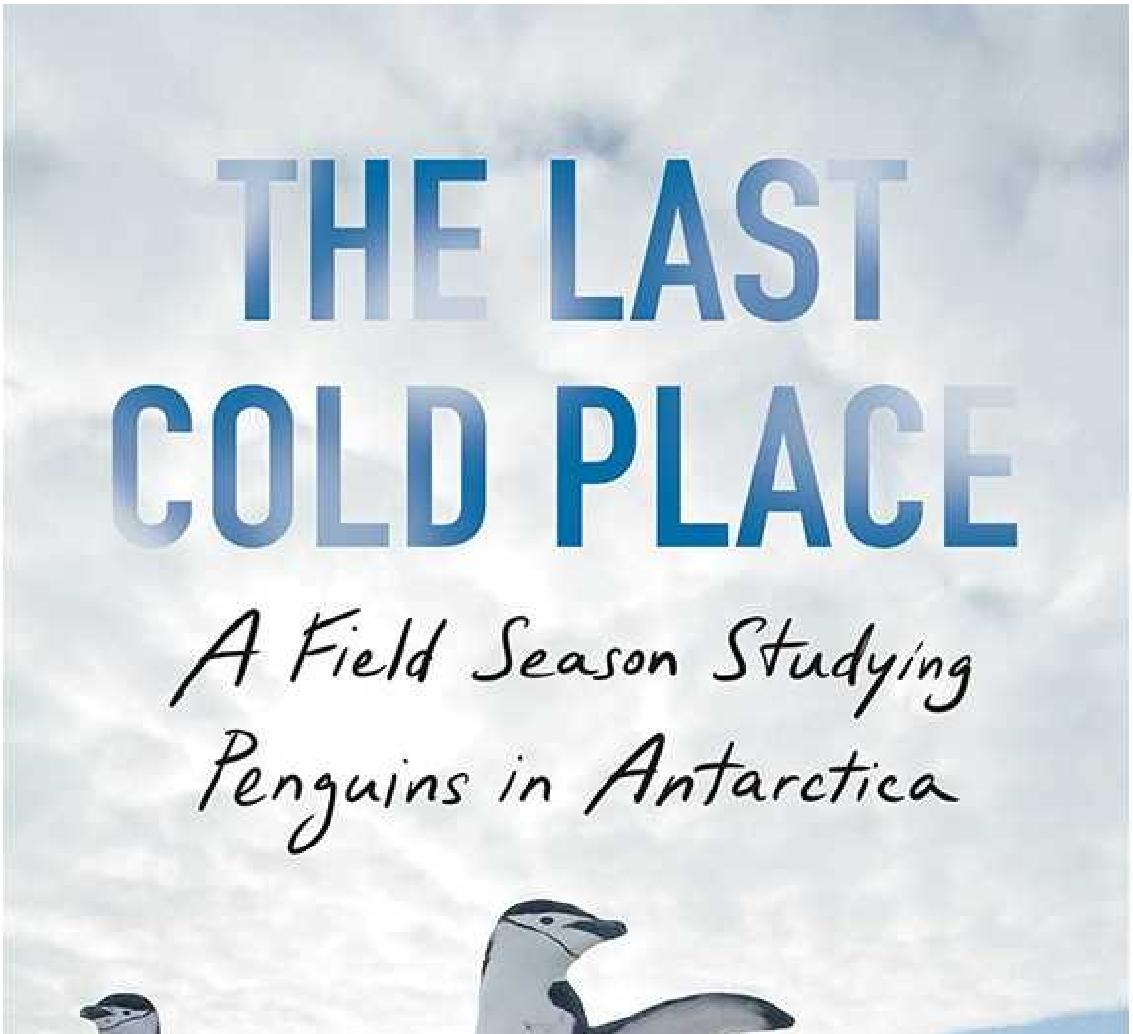
This article was downloaded by **calibre** from https://www.nature.com/articles/d41586-023-01938-6

- BOOK REVIEW
- 14 April 2023

Why incentives work — or don’t, and more: Books in brief

Andrew Robinson reviews five of the best science picks.

- Andrew Robinson


THE LAST
COLD PLACE
A Field Season Studying
Penguins in Antarctica
NAIRA DE GRACIA

The Last Cold Place

Naira de Gracia *Scribner* (2023)

The Western Antarctic Peninsula is warming five times faster than the global rate, with average winter temperatures 7 °C higher today than in 1950. But it remains a challenging environment for scientists, explains wildlife technician and conservation biologist Naira de Gracia in her exuberant memoir of five months studying penguins there in 2016–17. For each statistic about Antarctic marine species presented to policymakers, she says, there is “a grimy fieldworker like me”, covered in penguin muck and “smelling like fermented shrimp”.

HOW INCENTIVES REALLY WORK

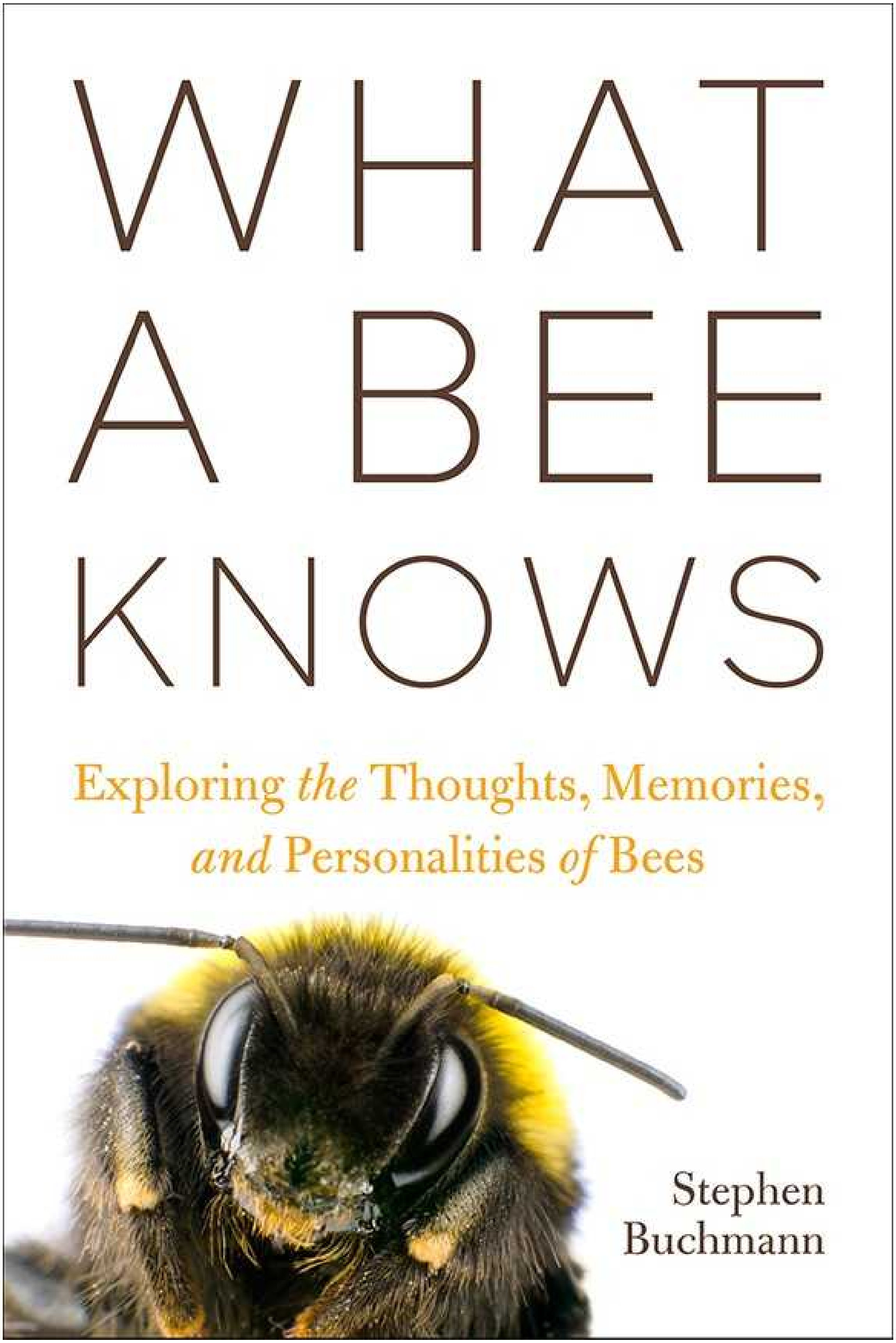
Mixed Signals

URI GNEEZY

Mixed Signals

Uri Gneezy *Yale Univ. Press* (2023)

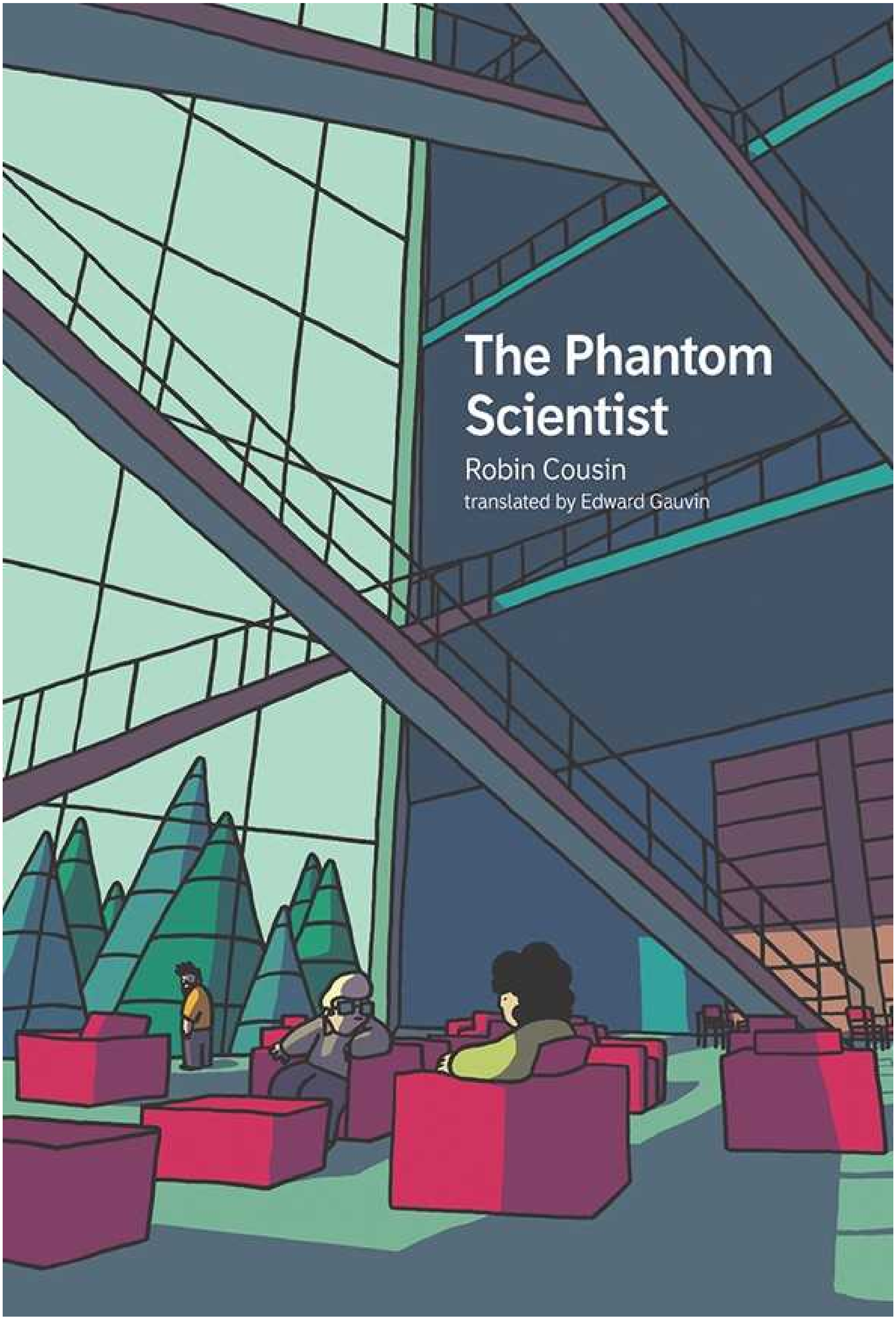
Incentives come in many guises. In Ohio in 2021, people were bribed to get COVID-19 vaccinations with tickets for a US$1-million lottery. In New Jersey, the prize was dinner with the governor. Behavioural economist Uri Gneezy thinks the latter “better and smarter”, despite attracting less press, because it signalled that the governor cared about vaccination. Other “mixed signals” in this stimulating book range from family life to politics. Gneezy argues that for an incentive to be effective, its signal must align with the sender’s goal.


WHAT
A BEE
KNOWS
Exploring the Thoughts, Memories, and Personalities of Bees
Stephen Buchmann

What a Bee Knows

Stephen Buchmann *Island* (2023)

Bee books tend to be full of intriguing facts and tantalizing possibilities. This rigorous but delightful work by pollination ecologist Stephen Buchmann, who has studied the insects for half a century, is no exception. The world has some 21,000 distinct bee species; a bee can pollinate as many as 10,000 flowers daily; and its brain houses only one million neurons, compared with at least 80 billion neurons in the human brain. Yet bees are "self-aware, they're sentient, and they possibly have a primitive form of consciousness", he writes.

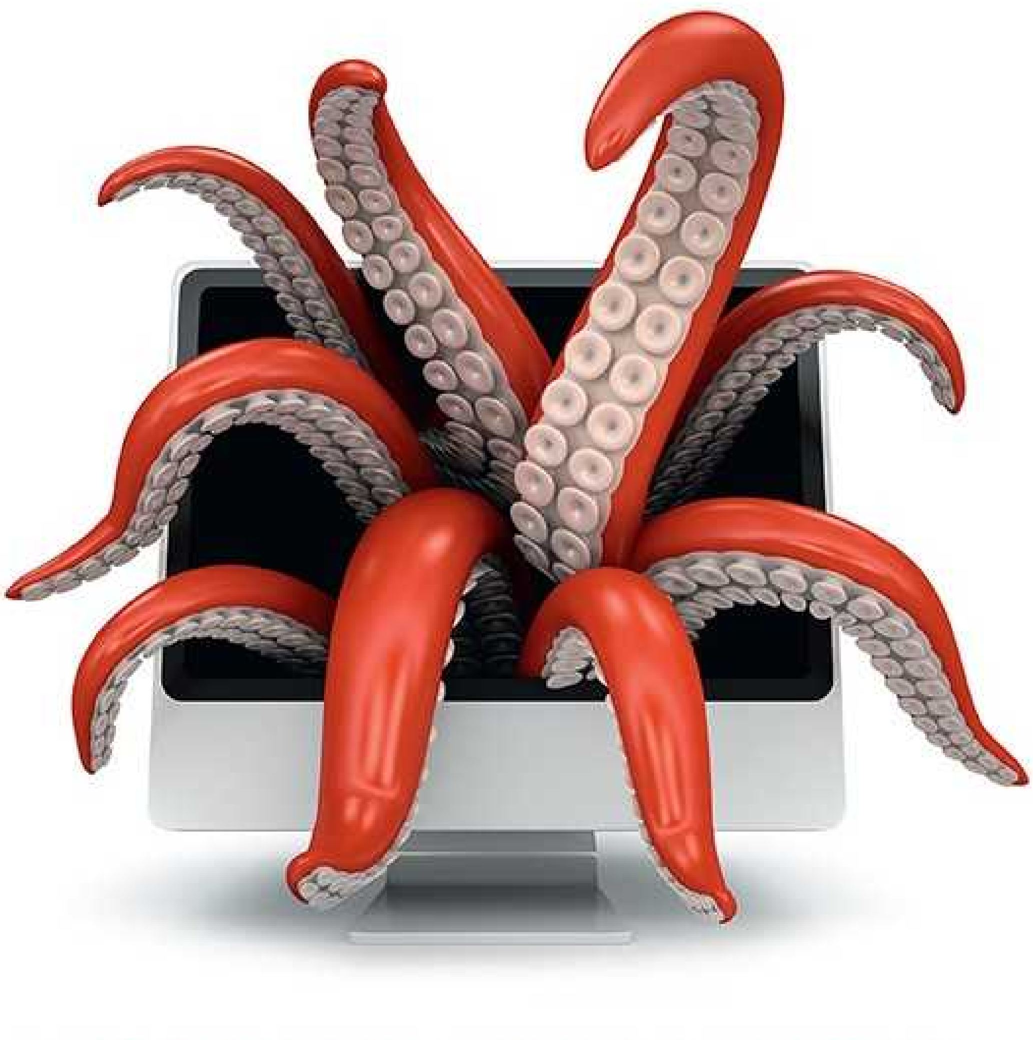

The Phantom Scientist
Robin Cousin
translated by Edward Gauvin

The Phantom Scientist

Robin Cousin (Transl. Edward Gauvin) *MIT Press* (2023)

In 2021, MIT Press published *The Curie Society*, an action-adventure graphic novel with broad appeal. Its latest graphic novel, visualized by Robin Cousin and translated from his French text, is much more specialized. Informed by work done at the French national research centre CNRS in Paris, it dramatizes a major unsolved challenge in computer science, known as P vs NP. Set in a fictional institute in a forest, from which a mysterious researcher suddenly vanishes, it is part thriller, part mystery, part systems theory.

THE INTERNET OF ANIMALS



DEBORAH LUPTON

The Internet of Animals

Deborah Lupton *Polity* (2023)

This academic study by sociologist Deborah Lupton analyses how the Internet depicts and changes relationships between humans and animals, including pets. The book examines how animals across a variety of species and in many spaces are incorporated into diverse digital technologies, and how this has affected human attitudes to animals. Inevitably, says Lupton, digitization and datafication have sometimes promoted “attitudes and practices contributing to animal mistreatment and cruelty”.

Nature **618**, 456 (2023)

doi: https://doi.org/10.1038/d41586-023-01269-6

This article was downloaded by **calibre** from https://www.nature.com/articles/d41586-023-01269-6

Opinion

| Next section | Main menu | Previous section |

Work

- **A family affair: how scientist parents' career paths can influence children's choices** [12 Jun 2023]
  • Early exposure to science inspired four researchers to pursue research careers.
- **How I use science to protect my people's birthright** [12 Jun 2023]
  • Jean-Luc Kanapé combines the ancestral knowledge of his Indigenous Canadian community, the Innu, with technology to protect the region's caribou from predators and environmental damage.

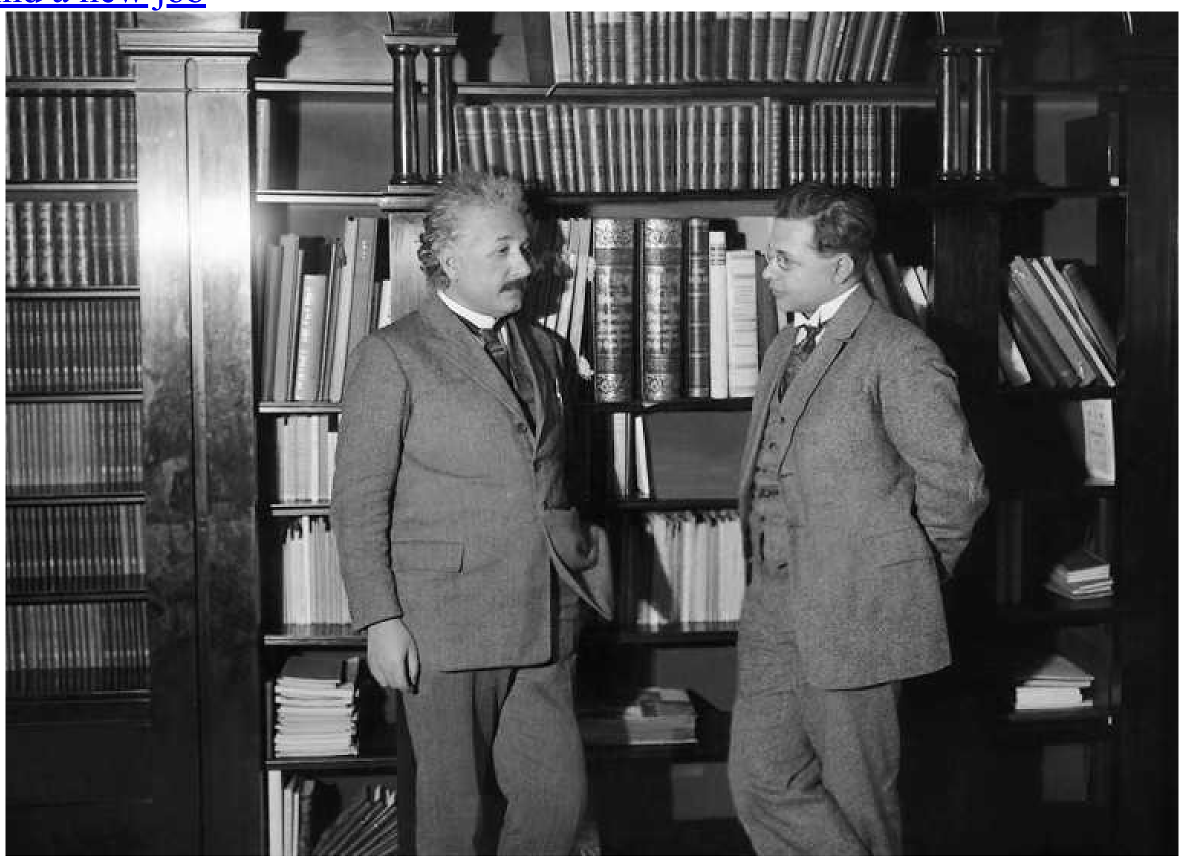
- CAREER FEATURE
- 12 June 2023

A family affair: how scientist parents' career paths can influence children's choices

Early exposure to science inspired four researchers to pursue research careers.

- Andy Tay

Find a new job



Albert Einstein and his son Hans-Albert both pursued distinguished careers in science.Credit: Süddeutsche Zeitung Photo/Alamy

Children whose parents have science degrees are twice as likely to pursue science degrees themselves than are those whose parents have degrees in other fields (N. Tilbrook and D. Shifrer *Soc. Sci. Res.* **103**, 102654; 2022). Scientist parents can be role models for their children and often provide early exposure through science-focused extracurricular activities. Their children can see at first-hand the highs and lows of a career in academic and industry research — the discoveries, collaborations and opportunities to live and work abroad. But that can be tempered by intense workloads, temporary contracts, pressures to publish and time away from families.

Four researchers share how their parents influenced their choice of a research career and how their own parenthoods have influenced their science.

FRED CHANG: Respect personal choices and decisions

Professor of cell and tissue biology at the University of California, San Francisco.

My parents immigrated from Taiwan to the United States in the 1950s to pursue graduate studies in engineering. My father, David Chang was a mechanical engineer who started a company in our garage, so my childhood was surrounded by electrical machinery and tools. My mother, Helen Chang, worked as a staff scientist at a diabetes lab at Stanford University in California. She introduced me to the environment of a biomedical lab and trained me to work in one. My parents placed a high priority on getting me the best education possible and gave me opportunities to broaden my education in maths and science.

In my early 30s, I married and had two children. I am a cell biologist and my ex-wife is a professional musician, so my daughter and son grew up with both music and science at home. They spent many formative summers with me at Woods Hole in Cape Cod, where I work as a summer investigator at

the Marine Biological Laboratory. Woods Hole is like a summer camp for scientists, and my children got to see how much fun I had making discoveries while collaborating with friends and colleagues.

Woods Hole also operates a science school at which my children learnt how to observe and explore the rich natural environments at the seashore. They're now in their late twenties. My daughter has always been fascinated by the history of Earth, and she's now a geologist. My son is a mechanical engineer who enjoys the practicality of building structures.

In my 40s, I came out as a gay man. It was an extraordinarily difficult process that took many years; I regard my coming out as my most courageous act. Although this was a challenging time for everyone in the family, we gradually adapted to the changes. My children have been an important source of support, and they fully support me and my partner. I would like to think that seeing me navigate my identity has had a positive influence on my children. Both have grown to be empathic and respectful humans.



Fred Chang (left) with his parents and two children.Credit: Fred Chang

LOTTE DE WINDE: Learn to compartmentalize and prioritize

Research associate at Amsterdam UMC location VU in the Netherlands.

My father is Han de Winde, a biotechnology researcher at Leiden University in the Netherlands. My mum trained as a paediatric nurse and has been working for almost 25 years as a nurse practitioner. Her name is Marga de Winde-van Zijl. When I was born, my dad was pursuing his PhD, but even after he became a professor, he did not miss any important moment of my life. He has shown me that it is possible to balance work and life well and how compartmentalization can help to achieve that. During my school holidays, I used to join my dad at work. I used to refill his pipette-tip boxes,

for example, and I enjoyed being in the lab environment. Later, he took me to open days at various Dutch universities, where we participated in medicine and science-related programmes and activities.

I initially wanted to become a physician, but as an undergraduate, I was deeply drawn to the study of our immune system. I wanted to understand why a system that is made to keep us healthy was failing to eradicate cancer. Now, I study lymphoma. My father read my applications to graduate school and gave me advice on how to strengthen my personal statements to show my interest in research. My parents also encouraged me to try my hand at many things. As a parent of a 1.5-year-old daughter, I want to be able to do the same for her.



[Are you a postdoc working in academia or industry? Share your career experiences with Nature]

I will fully support my daughter if she chooses a similar career, because research can have a positive impact on society and it's great for someone with a curious mind. But most if all, I want her to do something that makes her happy, whether in science or other fields.

Deciding to start a family was not an easy decision. My partner and I have been together since 2009 and moved to the United Kingdom in 2017. We decided to start a family only after returning to the Netherlands in 2020. We

felt that we had better job security there, and were closer to our families. There is also a commonly accepted practice for new parents to work four days in the office or the lab in the Netherlands, so that they can spend more time with their children in the early years of their life. These conditions gave us the confidence to start a family.

Becoming a parent has taught me a few valuable lessons that have benefited my work. I have learnt to compartmentalize my roles at work and home. I used to feel guilty when I was at work, because I could not take care of my child, yet I also felt that I was not giving enough time to my research. After I decided to give 100% to my research when in the lab and 100% to my child when at home, it improved my work productivity and the quality of my family time.

MARK PRAUSNITZ: Parenthood has parallels with professorship

Regent's professor of chemical and biomolecular engineering at Georgia Institute of Technology in Atlanta.

I grew up near the University of California, Berkeley, where my father, John Prausnitz, a chemical engineer, is now an emeritus professor. My late mother, Susan Prausnitz, was a paralegal. Many friends in their social circle are also researchers. Living in that environment gave my elder sister and me an early glimpse of what a career in science would be like. This influence was through soft power that focused on people and the joys of science rather than on hard technical content discussed over dinner or used as a lens for interpreting the world. My sister became a health-care researcher and I decided to follow in my father's footsteps and become a chemical engineer.



[Juggling research and family life: honest reflections from scientist dads](#)

A lesson I learnt from my father is epitomized by his lecture entitled '[Chemical engineering and the other humanities](#)', which he gave several times in the 1990s. Although I was already a young professor when he gave this particular lecture, he has been conveying the messages in it to me in direct and indirect ways ever since I was a child. In particular, he explains why scientific research is ultimately a human endeavour that affects society and how society in turn affects science. This perspective has influenced the research I do, which is to adapt engineering technologies to improve drug delivery and other medical interventions through simple, low-cost solutions that improve patient access.

My training as an engineer influences my mentorship style at work as well as my parenting at home. Engineering often emphasizes efficiency and teamwork to complete large projects, and this approach influences how I run my lab. I seek to prioritize activities that require my involvement and delegate others between the 26 members of my research group. This approach has spilled over into my home life with my wife — public-health professional Cindy Weinbaum — and three children. My wife and I needed to identify which activities we would prioritize doing with our kids, and which ones we might delegate, such as shuttling them to and from after-class activities when they were young.

Similarly, being a parent has taught me to be a better researcher. One of the great parallels between parenthood and being a professor is mentorship. I run my lab as a mentor, not a boss. I guide my students and postdocs in their research, offering suggestions (with varying levels of urgency) and helping them to become independent researchers. This mentorship style is also reflected in my parenting, and I see myself guiding my children to independence, too. It's a great feeling to see my children and my lab members grow and go on to make their own impacts on humanity.



Valerie Yang Shiwen with her partner, Alexander Yap, and son.Credit: Yan Jiejun

VALERIE YANG SHIWEN: Be disciplined and choose what's right for you

Assistant professor at the National Cancer Centre Singapore.

I remember a story from a colleague who was throwing a retirement party for a prominent professor. When invited to join, his children said that they didn't want to attend the party because their father had dedicated so much of his time to work that they did not feel that he was actually a father to them. This incident left a deep impression on me, and it reminds me not to further my career at the expense of my family.



The parenting penalties faced by scientist mothers

Both my father, Joseph Yang, and my mother, Theresa Yap, were general practitioners, so becoming a physician was a natural career decision. However, my dad would often encourage me to go into scientific research, telling me tales of the limitations of medical practice. For instance, he described how he would experiment with a combination of different off-the-shelf creams to achieve the best results for his patients with recalcitrant eczema, yet be unable to decipher why some patients fared better than others. Eventually, I split the difference and started studying for a PhD in oncology at the University of Cambridge, UK, in 2006.

I had my son in 2016, during my clinical specialty training, and received my first independent grant the day before I gave birth, so had to deliver both the baby and the research. After three months of maternity leave, I went back to work, and I saw myself missing some of my son's important milestones, such as sitting up independently, rolling, babbling and trying different foods for the first time. I was leaving for work before 5 a.m. and not returning until past 11 p.m., and often had to stay in hospital overnight on-call for either ward or intensive care unit coverage. So I decided to take six months of unpaid leave so that I could spend quality time with my son. It certainly felt like I was jeopardizing my career by delaying the exit exam for my clinical specialty, but I now know that I made the right choice. I cannot not turn back the clock to witness my son's milestones that I would otherwise have missed, but there would always be other grants and opportunities for me to grow my career.

Parenthood has taught me to be more disciplined in my work and to dedicate my time and limited resources to projects that really matter the most to me.

Nature **618**, 635-637 (2023)

doi: https://doi.org/10.1038/d41586-023-01939-5

These interviews have been edited for length and clarity.

This article was downloaded by **calibre** from https://www.nature.com/articles/d41586-023-01939-5

- WHERE I WORK
- 12 June 2023

How I use science to protect my people’s birthright

Jean-Luc Kanapé combines the ancestral knowledge of his Indigenous Canadian community, the Innu, with technology to protect the region’s caribou from predators and environmental damage.

- Patricia Maia Noronha



Jean-Luc Kanapé is Guardian of the Pipmuakan forest in Quebec, Canada. Credit: Ed Jones/AFP via Getty

I'm Guardian of the Pipmuakan forest in the Nitassinan territory of Quebec, Canada. I act as the eyes and ears of my people, the Innu, so that we can protect the caribou (*atik* in the Innu language). The caribou, or reindeer (*Rangifer tarandus*), is sacred to the Innu. For centuries, it provided us with food, clothes and tools.

In my work, I combine ancestral knowledge with new technologies. The knowledge includes how to assess the ice thickness and snow depth, and how to identify the freshness of animal tracks. Drones, satellite images and GPS are among the new technologies I use.

I travel across the Pipmuakan area at least four days a week on either my quad bike or snowmobile, looking for caribou tracks and other signs of herds and their predators. I also set traps for wolves.

Caribou are very sensitive, so I always avoid getting close to them. When I see new caribou tracks, I use the drone to identify their location and confirm the number of animals in the herd, whether they are male, female or both, and if there are new calves. The drone also allows me to spot herds behind dense thickets without scaring them.

Logging is the main threat to this forest and to the caribou. In winter, caribou eat lichen from trees. But logging destroys the lichen, and allows wolves — the caribou's main predator — to see their prey from afar.

In this picture from the third quarter of last year, I was crossing Kakuskanus lake in a pontoon boat to reach a remote area of the forest. My grandparents drowned in this lake when their canoe overturned. For 36 years, I couldn't visit this place. I was mourning.

But in 2017 I felt an urge to protect the land and came back. A century ago, there were thousands of caribou here. Now there are fewer than 200. Without the caribou, the Innu wouldn't have survived in this forest. Now it's time for us to help the caribou. I want my grandchildren and the following generations to experience the richness of this forest.

Nature **618**, 640 (2023)

doi: https://doi.org/10.1038/d41586-023-01941-x

This article was downloaded by **calibre** from https://www.nature.com/articles/d41586-023-01941-x

| Section menu | Main menu |

Research

- **Detection of phosphates originating from Enceladus's ocean** [14 Jun 2023]
  • Cassini's Cosmic Dust Analyzer mass spectra of ice grains emitted by Enceladus show the presence of sodium phosphates, suggesting that phosphorus is readily available in Enceladus's ocean in the form of orthophosphates.
- **Evidence for the utility of quantum computing before fault tolerance** [14 Jun 2023]
  • Experiments on a noisy 127-qubit superconducting quantum processor report the accurate measurement of expectation values beyond the reach of current brute-force classical computation, demonstrating evidence for the utility of quantum computing before fault tolerance.
- **Hadaean to Palaeoarchaean stagnant-lid tectonics revealed by zircon magnetism** [14 Jun 2023]
  • Magnetic palaeointensity data from the Barberton Greenstone Belt (South Africa) as well as the Jack Hills (Western Australia) show nearly constant palaeofield values between 3.9 Ga and 3.4 Ga, providing evidence for stagnant-lid mantle convection.
- **Unveiling the transition from niche to dispersal assembly in ecology** [07 Jun 2023]
  • This study shows through a field experiment that tropical intertidal communities switch from a niche- to a dispersal-assembled regime as immigration rates rise.
- **A median fin derived from the lateral plate mesoderm and the origin of paired fins** [24 May 2023]
  • We identify that the larval zebrafish unpaired pre-anal fin fold is derived from the lateral plate mesoderm, can be readily duplicated, and thus may represent a developmental intermediate between median and paired fins.
- **Northwest African Neolithic initiated by migrants from Iberia and Levant** [07 Jun 2023]
  • Genome sequencing of nine individuals shows ancestry shifts in the Neolithization of northwestern Africa that probably mirrored a heterogeneous economic and cultural landscape in a more multifaceted process than observed in other regions.
- **Geometric constraints on human brain function** [31 May 2023]

• Cortical and subcortical activity can be parsimoniously understood as resulting from excitations of fundamental, resonant modes of the brain's geometry rather than from modes of complex interregional connectivity.

- **The effects of cash transfers on adult and child mortality in low- and middle-income countries** [31 May 2023]
  • A survey of mortality data from low- and middle-income countries shows that government-led cash transfer programmes are associated with decreased mortality among women and children under five years of age.
- **A pan-influenza antibody inhibiting neuraminidase via receptor mimicry** [31 May 2023]
  • The neuraminidase-targeting monoclonal antibody FNI9 potently inhibits the enzymatic activity of influenza A and B viruses via receptor mimicry.
- **Axonemal structures reveal mechanoregulatory and disease mechanisms** [31 May 2023]
  • Detailed atomic models of axonemes from algal flagella and human respiratory cilia, which are hair-like protrusions from cells that enable motility and clear mucus from human airways, could provide insights into how they function.

- Article
- Open Access
- Published: 14 June 2023

Detection of phosphates originating from Enceladus's ocean

- Frank Postberg ,
- Yasuhito Sekine,
- Fabian Klenner,
- Christopher R. Glein,
- Zenghui Zou,
- Bernd Abel,
- Kento Furuya,
- Jon K. Hillier,
- Nozair Khawaja,
- Sascha Kempf,
- Lenz Noelle,
- Takuya Saito,
- Juergen Schmidt,
- Takazo Shibuya,
- Ralf Srama &
- …
- Shuya Tan

Show authors

Nature **volume 618**, pages 489–493 (2023)

- 14k Accesses
- 2185 Altmetric
- Metrics details

Abstract

Saturn's moon Enceladus harbours a global[1] ice-covered water ocean[2,3]. The Cassini spacecraft investigated the composition of the ocean by analysis of material ejected into space by the moon's cryovolcanic plume[4,5,6,7,8,9]. The analysis of salt-rich ice grains by Cassini's Cosmic Dust Analyzer[10] enabled inference of major solutes in the ocean water (Na^+, K^+, Cl^-, HCO_3^-, CO_3^{2-}) and its alkaline pH[3,11]. Phosphorus, the least abundant of the bio-essential elements[12,13,14], has not yet been detected in an ocean beyond Earth. Earlier geochemical modelling studies suggest that phosphate might be scarce in the ocean of Enceladus and other icy ocean worlds[15,16]. However, more recent modelling of mineral solubilities in Enceladus's ocean indicates that phosphate could be relatively abundant[17]. Here we present Cassini's Cosmic Dust Analyzer mass spectra of ice grains emitted by Enceladus that show the presence of sodium phosphates. Our observational results, together with laboratory analogue experiments, suggest that phosphorus is readily available in Enceladus's ocean in the form of orthophosphates, with phosphorus concentrations at least 100-fold higher in the moon's plume-forming ocean waters than in Earth's oceans. Furthermore, geochemical experiments and modelling demonstrate that such high phosphate abundances could be achieved in Enceladus and possibly in other icy ocean worlds beyond the primordial CO_2 snowline, either at the cold seafloor or in hydrothermal environments with moderate temperatures. In both cases the main driver is probably the higher solubility of calcium phosphate minerals compared with calcium carbonate in moderately alkaline solutions rich in carbonate or bicarbonate ions.

Main

Enceladus's global ocean lies under an ice crust and above a rocky core where tidal dissipation[18,19] is suspected to drive hydrothermal activity[4,5]. There are several lines of evidence describing how volatile[5] and dissolved materials[3,4,11] from the rocky core are either emitted by the plume in the gas phase or incorporated into ice particles, respectively.

Cassini's Cosmic Dust Analyzer (CDA) recorded time-of-flight (ToF) mass spectra with a mass resolution of $m/\Delta m \approx 10–50$ (refs. [10,20]) for cations generated by high-velocity impacts of individual grains onto the instrument's rhodium target. The E-ring of Saturn is formed by ice grains escaping Enceladus's plume into orbits around Saturn[21], and hence the analysis of these grains by CDA provides important insights into the composition of the subsurface ocean—including a rich variety of organic compounds[6,7,8,9]—with much better statistics compared with data from the rare occasions when Cassini traversed the plume itself.

In a previous analysis it was shown that a fraction of these ice grains, called Type 3 particles, contain salts at significantly higher concentrations than all other ice grains in the plume and the E-ring[3,11]. By co-addition of spectra from 107 individual detections in the E-ring, average salt concentrations of 0.5–2.0% by weight (around 0.07–0.30 M) have been deduced for Type 3 grains, with NaCl, $NaHCO_3$, Na_2CO_3 and KCl as the most abundant constituents. Because it was inferred that these grains are derived from aerosolized ocean water[3], the salt chemistry of the plume and E-ring grains would reflect that of the ocean.

Phosphorus is an element essential for planetary habitability[12,13,14], but to date it has not been detected in an ocean beyond Earth. Previous geochemical modelling suggested that phosphate might be scarce in the ocean of Enceladus and other icy ocean worlds[15,16]. However, more recent modelling of mineral solubilities in Enceladus's ocean suggests that phosphate could be relatively abundant (roughly 10^{-7}–10^{-2} M), depending on the compositional characteristics of the ocean such as its pH and carbonate content, as well as on attainment of certain chemical equilibria between ocean water and seafloor alteration minerals[17].

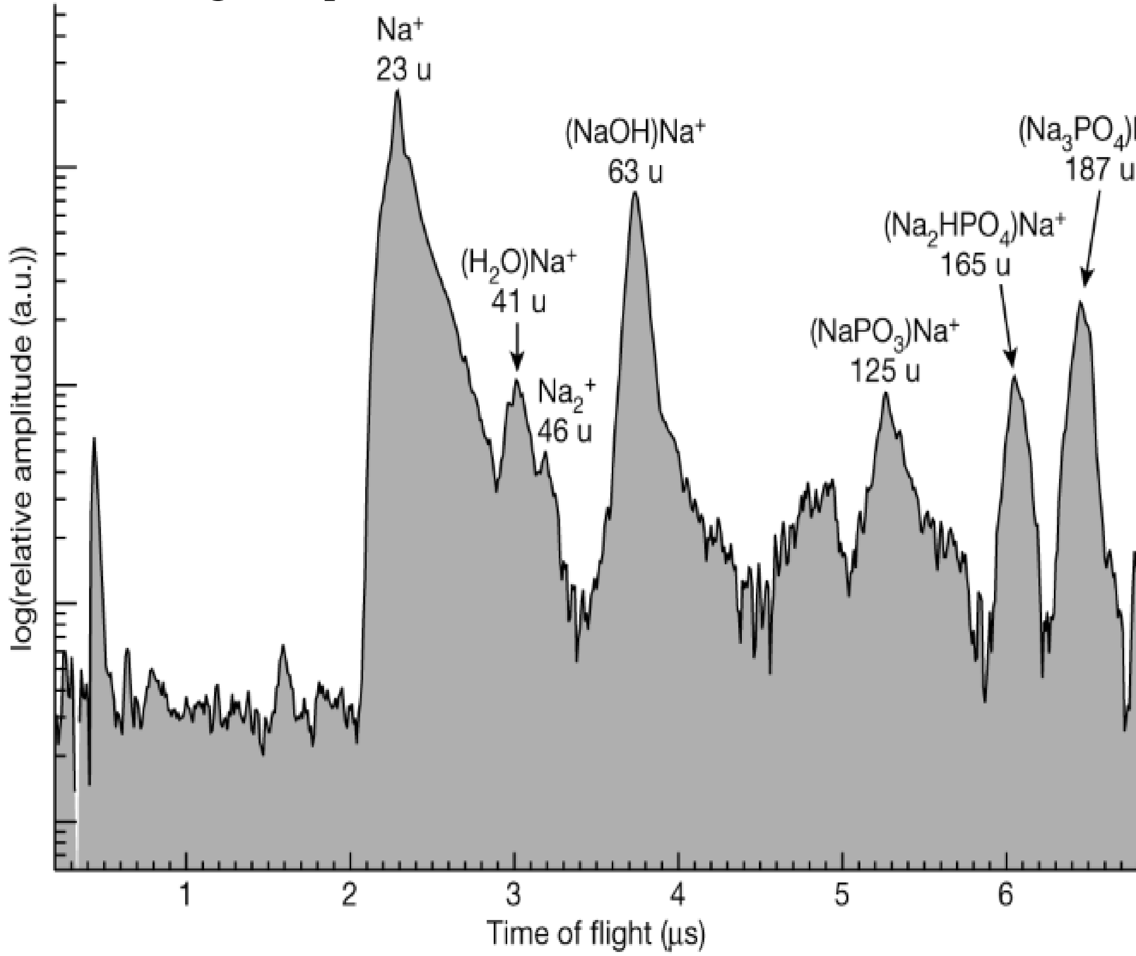
In this work we present CDA mass spectra of a population of E-ring ice grains that show the presence of sodium phosphates. We then perform laboratory analogue experiments to quantitatively establish that Enceladus's ocean is rich in dissolved phosphate. Lastly, we show through water–rock alteration experiments and complementary chemical speciation calculations that phosphate-rich fluids are an inevitable outcome of alkaline and

carbonate-rich ocean water reacting with unaltered carbonaceous chondritic rock.

Results

We conducted a comprehensive survey of 345 Type 3 particles (Methods, Extended Data Table 1). Our analysis led to the detection of a rare subtype (nine particles; Methods, Extended Data Table 2) whose spectra show highly significant features of phosphate species: in addition to the defining mass spectral peaks of Na-salt-rich water ice grains at 23 and 63 u (ref. [3]), there were peaks corresponding to molecular masses of 125, 165 and 187 u forming a pattern we found to be unique to sodium phosphates at high concentrations within these ice grains (Fig. 1 and Methods).

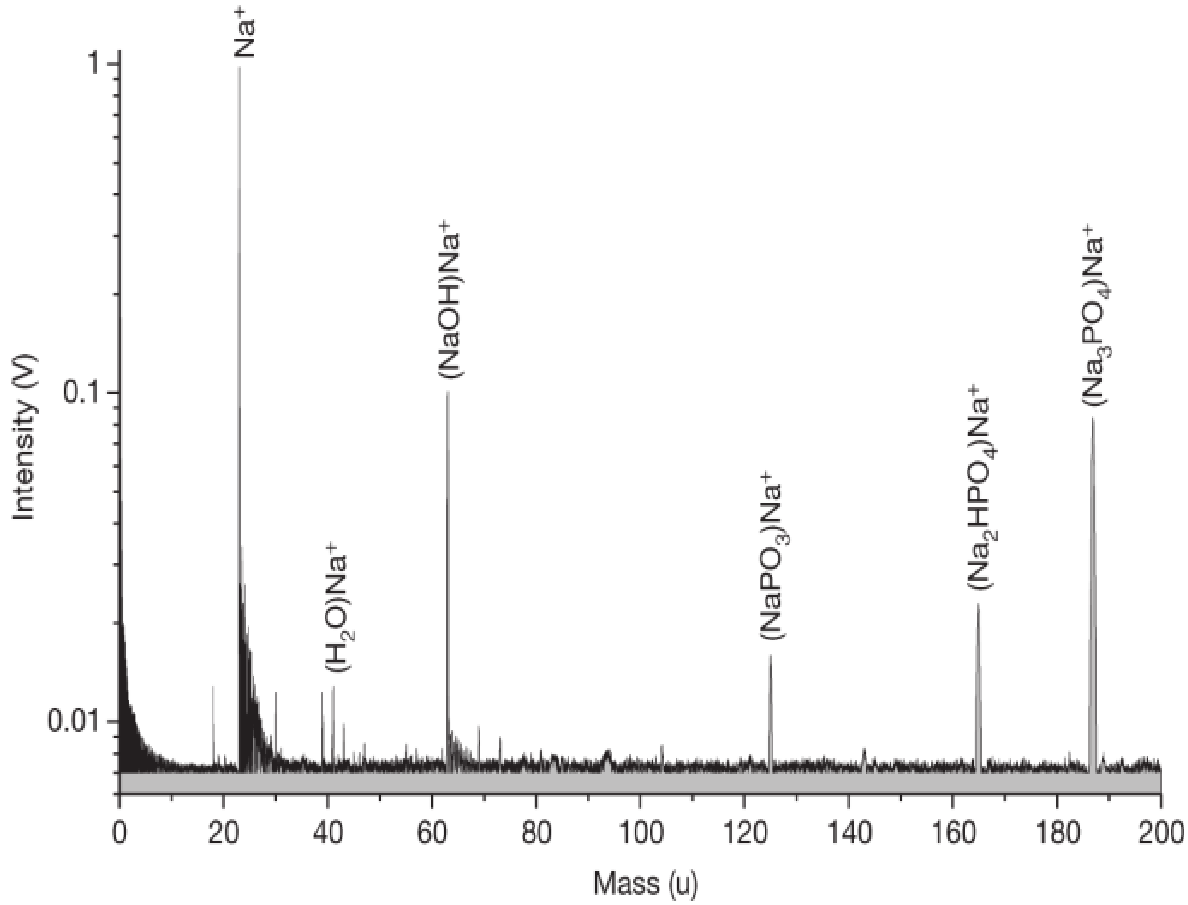
Fig. 1: CDA cation spectrum co-added from nine baseline-corrected individual ice grain spectra.



The mass lines signifying a high-salinity Type 3 spectrum are Na^+ (23 u) and $(NaOH)Na^+$ (63 u) with secondary Na-rich signatures of $(H_2O)Na^+$ (41 u) and Na_2^+ (46 u). Sodium phosphates are represented by phosphate-bearing Na-cluster cations, with $(Na_3PO_4)Na^+$ (187 u) possessing the highest amplitude in each spectrum followed by $(Na_2HPO_4)Na^+$ (165 u) and $(NaPO_3)Na^+$ (125 u). The first two unlabelled peaks at the beginning of the spectrum are H^+ and C^+, stemming from target contamination[39]. a.u., arbitrary units.

To obtain a quantitative assessment of the composition of these phosphate-bearing ice grains, an analogue experiment was used to simulate the impact ionization of ice grains via a laser pulse targeting a micrometre-sized water beam[22]. This technique, laser-induced desorption of ions and ionic aggregates[23], has previously reproduced accurate CDA ice grain spectra[3,9]. Despite the higher amplitude of the $(Na_3PO_4)Na^+$ peak compared with that of $(Na_2HPO_4)Na^+$, the peak pattern can best be reproduced by dissolving Na_3PO_4 and Na_2HPO_4 at an inverse salt molar mixing ratio between the two salts of 1:2.5–1:25 (Fig. 2), indicating the dominance of monohydrogen phosphate (Na_2HPO_4) over phosphate (Na_3PO_4) in these ice grains. The dominance of Na_2HPO_4 is substantiated by the presence of $(NaPO_3)Na^+$ (125 u), a charged molecular meta phosphate aggregate (Methods, Extended Data Fig. 4) that formed from Na_2HPO_4 in our analogue experiments, but not from Na_3PO_4 alone (Methods, Extended Data Figs. 1 and 2). With the molar ratio of Na_2HPO_4:Na_3PO_4 kept above 2.5:1.0, a concentration ranging from 0.05 to 0.60 M of total phosphate salts then gave a close match with CDA spectra of phosphate-rich grains (Methods).

Fig. 2: Spectrum from the LILBID analogue experiment reproducing the features in the CDA spectrum.



An aqueous solution of 0.420 M Na_2HPO_4 and 0.038 M Na_3PO_4 was used. All major characteristics of the CDA spectrum of phosphate-rich grains (Fig. 1) are reproduced at the higher mass resolution of the laboratory mass spectrometer (roughly 700 $m/\Delta m$). Note: this solution is not equivalent to the inferred ocean concentration. To derive the latter quantity, the concentration determined in these P-rich grains must be averaged over the entire dataset of salt-rich ice grains (see below).

We also tested for the presence of phosphites (for example, Na_2HPO_3, where P has an oxidation state of +III compared with +V in the orthophosphates Na_3PO_4 and Na_2HPO_4). Because the acid dissociation constant (pK_a) of $H_2PO_3^-$ is about 6.2 at 0 °C (ref. [24]), this species is unlikely to be relevant to Enceladus's alkaline ocean[3,11]. Raising the concentration of Na_2HPO_3 in the phosphate mixture to 5 mM and above led to a clear mass line at 149 u due to $(Na_2HPO_3)Na^+$ (Methods and Extended Data Fig. 3). However, this peak was not observed in CDA spectra (5.7–5.8 μs in Fig. 1), thus providing an

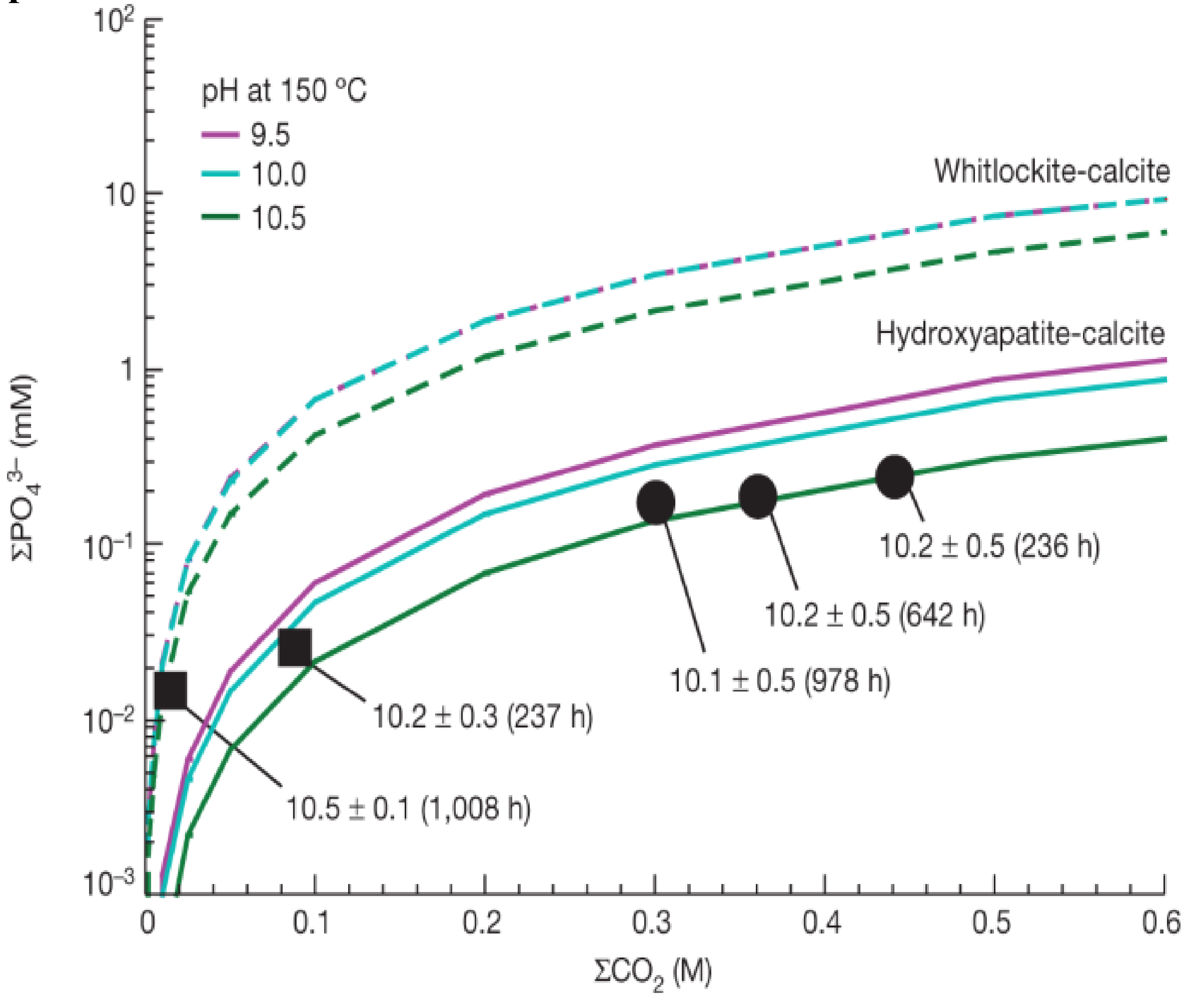
upper limit on phosphite concentrations in these ice grains that is one to two orders of magnitude below the required phosphate concentration.

We now follow the approach established in Postberg et al.[3], which assumes that the total ensemble of salt-rich grains emitted currently by the plume is the best representation of the composition of the source water. Thus, to infer the concentration in the ocean we calculate the average phosphate concentration in all Type 3 grains. We 'dilute' the concentration in the nine spectra of the new phosphate-bearing subtype (2.6 ± 0.9% by number) with the 336 spectra of other Type 3 grains—which are mostly dominated by NaCl or Na_2CO_3 (ref. [3])—and do not show phosphates at detectable concentrations ($[\Sigma PO_4^{3-}] < 5$ μM). Including the ±0.9% s.e.m. from counting statistics and the inferred concentrations of total dissolved phosphate ($[\Sigma PO_4^{3-}] = 0.05 – 0.60$ M; $[\Sigma PO_4^{3-}] = [PO_4^{3-}] + [HPO_4^{2-}] + [H_2PO_4^{-}]$, where [*X*] denotes the molar concentration of species *X*), this yields a phosphate concentration range of 0.8–21 mM for the Enceladean ocean. From our analogue experiments, the upper limit inferred from the nondetection of phosphites lies at least one order of magnitude below phosphate concentrations. This indicates a clear predominance of orthophosphates as the most abundant P-bearing species in the Enceladean ocean.

To interpret the CDA detection of phosphates we performed hydrothermal laboratory experiments with powdered samples from a carbonaceous chondrite (Jbilet Winselwan (CM-2 type))[25] as an analogue of Enceladus's core material[26]. The main purpose of the hydrothermal experiments was to elucidate the geochemical controls on phosphate solubility under conditions relevant to Enceladus. The chondritic material was exposed to two different Na-carbonate-enriched solutions (Run no. 1: initial $\Sigma CO_2 = 0.2$ M; Run no. 2: initial $\Sigma CO_2 = 1.0$ M), where the total dissolved inorganic carbon concentration is given by $\Sigma CO_2 = [CO_{2,aq}] + [HCO_3^{-}] + [CO_3^{2-}] + [NaCO_3^{-}] + [NaHCO_3]$. The starting solution also contained 0.5 M $\Sigma NH_3 = [NH_{3,aq}] + [NH_4^{+}]$, a plausible concentration for Enceladus's ocean[5]. An elevated temperature (150 °C) was used in the experiment to accelerate reaction rates. This allowed us to study key aspects of the water–rock interactions occurring in Enceladus at time scales accessible in the laboratory of up to

about 1,000 h. Fluid samples were collected periodically from the reaction cell during the experiments to determine pH and fluid composition, including the concentrations of P and ΣCO_2 (Methods). The initial $[\Sigma CO_2]$ adopted for Run no. 2 (1.0 M) was higher than the value inferred for the Enceladean ocean[27,28], but a large fraction of initial ΣCO_2 was removed during the experiments via formation of carbonate minerals (Methods), leading to more realistic levels of carbonate species (Fig. 3).

Fig. 3: Phosphate concentrations derived from hydrothermal experiments.



Experimental results (squares for Run no. 1 and circles for Run no. 2, see text) of P concentrations in hydrothermal fluids at 150 °C and 30 MPa are shown as a function of ΣCO_2 concentration. Numbers attached to data points are calculated in situ pH (Methods) and reaction times in hours. Solid and

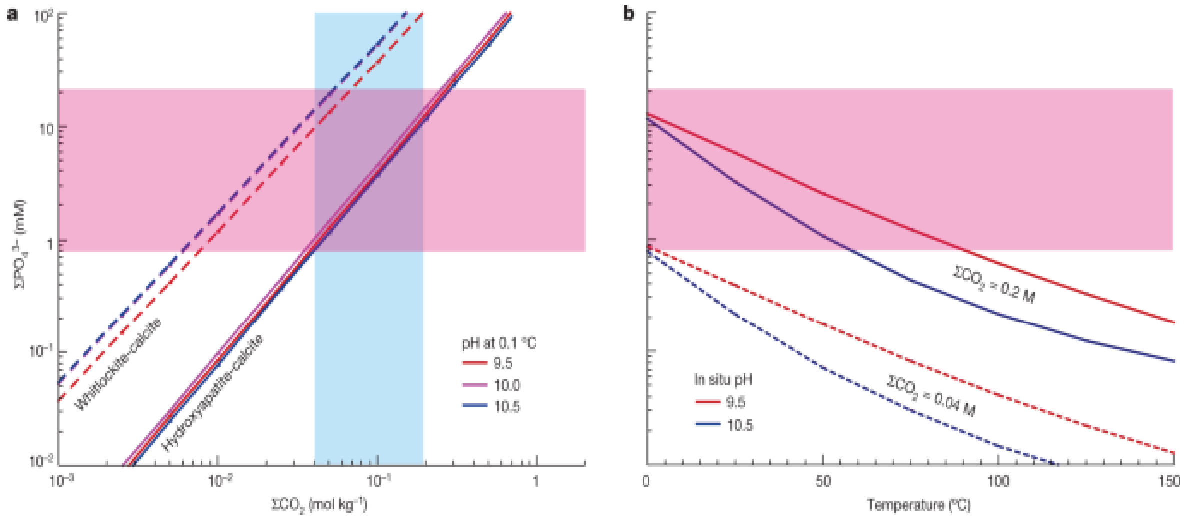
dashed curves represent results from thermodynamic equilibrium calculations (Methods) for hydroxyapatite-calcite and whitlockite-calcite buffer systems, respectively, evaluated at different pH values.

Our experiments show that the dissolved phosphate concentration increases with in situ measured ΣCO_2 (Fig. 3). This happens because an increase in CO_3^{2-} concentration (or, more correctly, aqueous activity) leads to dissolution of Ca phosphate minerals with the formation of stable Ca carbonates (that is, calcite)[17,29,30]. In comparison with thermodynamic equilibrium predictions, we find that the hydroxyapatite-calcite buffer system (hydroxyapatite + $5CO_3^{2-} + H^+ \leftrightarrow$ calcite + $3PO_4^{3-} + H_2O$) is generally applicable to our experimental results (Fig. 3). The relevance of this buffer is further supported by chemical analysis of the solid residues after Run no. 2, which demonstrates the presence of coexisting Ca phosphate with high Ca:P ratios (probably a mixture of hydroxyapatite and Ca-/Mg-whitlockite) and calcite (Methods, Extended Data Figs. 5–7, Extended Data Table 4). One data point in Fig. 3 ($\Sigma CO_2 = 0.01$ M in Run no. 1) lies significantly above the hydroxyapatite-calcite buffer, appearing to be more consistent with the whitlockite-calcite buffer (whitlockite (β-$Ca_3(PO_4)_2$) + $3CO_3^{2-} \leftrightarrow$ calcite + $2PO_4^{3-}$). We suspect that this change in behaviour is due to unusually high concentrations of Ca^{2+} and Mg^{2+} in Run no. 1 (0.8 and 0.3 mM, respectively; Methods, Extended Data Fig. 8, Extended Data Table 3), which may promote the precipitation of whitlockite[31]. This interpretation is supported by an electron probe microanalysis of the solid residue (Methods, Extended Data Table 4). Alternatively the CM chondrite, Jbilet Winselwan (Extended Data Figs. 9 and 10), used in our experiments is known to contain impact-induced, thermally metamorphosed materials[25], which could make some of the results less relevant to Enceladus. By contrast, buffer systems involving other phosphates (for example, Mg-whitlockite and merrillite) and carbonates (for example, dolomite) cannot explain the experimental results (Methods).

Figure 4a shows the modelled ΣPO_4^{3-} as a function of ΣCO_2 for water–rock interactions adopting the two buffer systems inferred from our high-temperature experiments, but now transferred to conditions corresponding to the cold seafloor of Enceladus (0.1 °C and 8 MPa). The P concentrations

determined from the CDA measurements in this work (ΣPO_4^{3-} = 0.8–21 mM), the independently derived constraint on ΣCO_2 (0.04–0.20 M)[27,28], as well as an alkaline pH (9.0–10.5)[3,4], are all fully consistent with these two buffer systems. The hydroxyapatite-calcite buffer is applicable for the entire range of ΣCO_2 (Fig. 4a), corresponding to ΣPO_4^{3-} of about 0.8–12 mM. The whitlockite-calcite buffer is applicable for low ΣCO_2 of about 0.04–0.07 M corresponding to ΣPO_4^{3-} in the range of about 7–21 mM at 0.1 °C.

Fig. 4: Comparison of observed and calculated concentrations of ΣPO_4^{3-} in fluids affected by water–rock reactions within Enceladus.



a, Relation between ΣPO_4^{3-} and ΣCO_2 at a temperature of 0.1 °C for the hydroxyapatite-calcite buffer system (solid lines) and the whitlockite-calcite buffer system (dashed lines). Constraints on ΣCO_2 obtained in previous studies[27,28] are indicated by the blue shaded area. The area highlighted in pink represents the range of ΣPO_4^{3-} constrained in this study from CDA data. **b**, Dependence of ΣPO_4^{3-} on temperature for the hydroxyapatite-calcite buffer and different values of pH and ΣCO_2. A similar diagram for the whitlockite-calcite buffer can be found in Extended Data Fig. 11.

Figure 4b shows the temperature dependence of ΣPO_4^{3-} in fluids controlled by the hydroxyapatite-calcite buffer. The effects of colder fluids, as used in the experiments, down to oceanic temperatures near 0 °C can be seen. The phosphate concentration increases with lower temperature at constant pH

because hydroxyapatite (and other reactants) becomes less stable relative to calcite (and other products) at lower temperatures. To achieve ΣPO_4^{3-} in the range 0.8–21 mM, as constrained by the CDA measurements, the temperature of the water–rock reactions that provide phosphate in Enceladus would need to be 80 °C or less if the pH is 9.5, and below 65 °C if the pH is 10.5 (Fig. 4b). Higher-temperature fluids could contain sufficient phosphate only if the fluid pH were lower. If the whitlockite-calcite buffer controls ΣPO_4^{3-} in Enceladus's ocean, the temperature could be as high as around 180 °C (Extended Data Fig. 11) for a maximum ΣCO_2 of 0.2 M. In the case in which high-temperature interactions with seafloor rocks provide the oceanic phosphate, its concentration in the fluids should be preserved on cooling to near subzero temperature within the cold ocean. This scenario would, however, be relevant only if the time scale of cycling the ocean through hydrothermal systems is shorter than that of low-temperature dissolution of phosphate minerals at the seafloor.

Given the consistency of the measured ΣPO_4^{3-}, ΣCO_2 and pH of Enceladus's ocean, we suggest that thermodynamic equilibrium between Ca phosphate (for example, hydroxyapatite) and calcite is probably achieved in both the cold global ocean and hydrothermal fluids of moderate temperature.

Discussion

The CDA detection of ice grains with high concentrations of orthophosphates indicates that phosphorus is readily available at the top of Enceladus's ocean (that is, the plume source region). Even with a conservative margin, our estimate indicates concentrations in the order of at least hundreds of micromolar, several 100-fold the average phosphate abundance in Earth's oceans[32]. The heterogenous distribution of salts in the emitted ice grains might suggest some level of heterogeneity in the ocean and that actual concentrations could be locally higher or lower than the bulk number calculated in this work. Alternatively, yet unknown mechanisms might separate different salts when ocean material is transported upwards in the vents. However, even without knowing the details of the subsurface processes, we assume that the total ensemble of salt-rich grains emitted by

the plume is a reasonable representation of the average ocean composition near its surface.

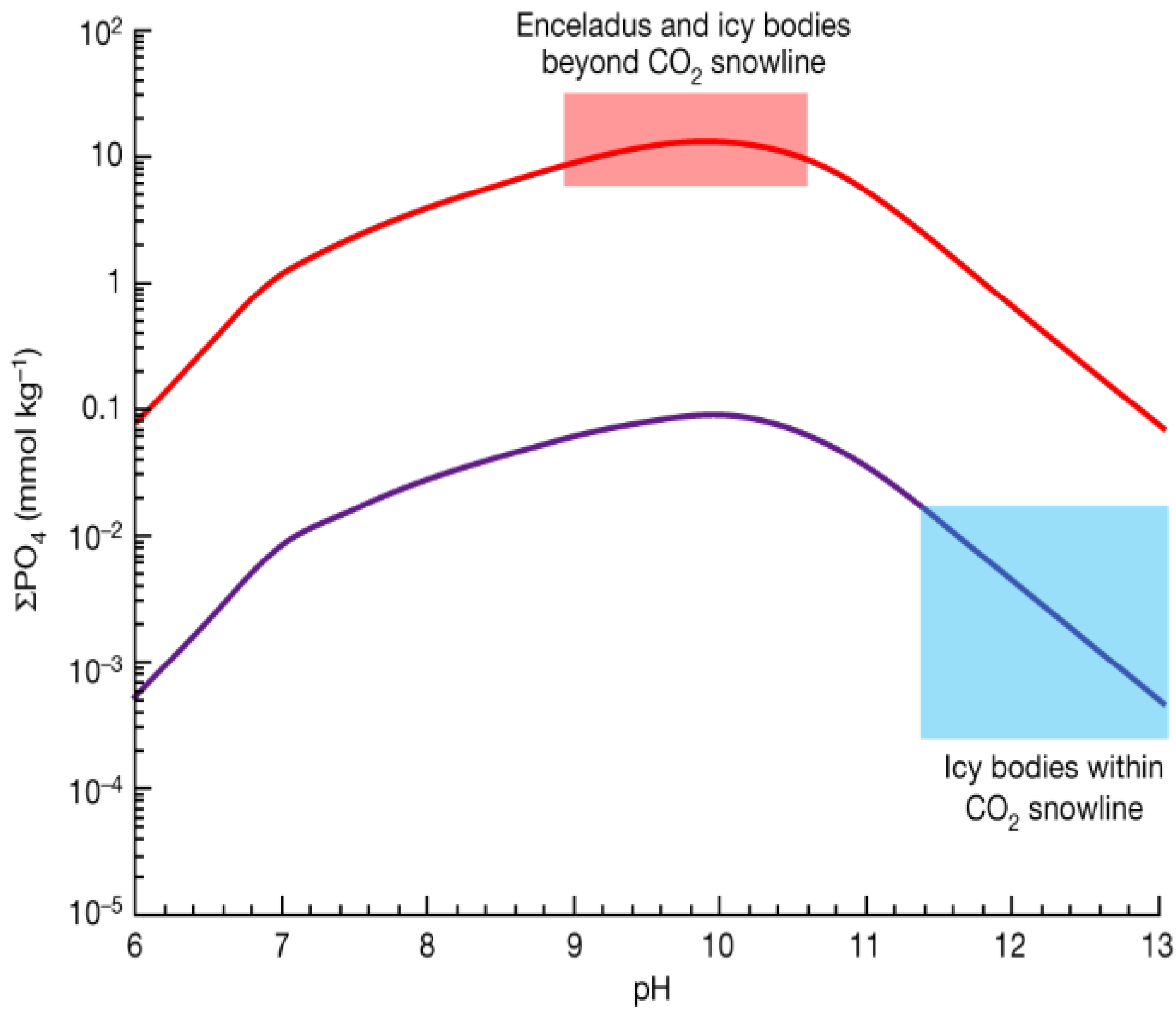
A phosphate:phosphite ratio of at least around 10:1, as implied by our results, is consistent with the much greater thermodynamic stability of phosphates relative to phosphites under Enceladean ocean conditions[17]. In addition, the nondetection of phosphites suggests that aqueous alteration of iron phosphides (the most probable accreted form of P) is no longer a significant source of P for Enceladus's ocean. This process would produce metastable phosphites[33]. The inferred lack of contemporary alteration of phosphides is consistent with the presence of a low-density rocky core that has already experienced extensive water–rock interaction[2,18,34,35]. Phosphites could have been produced when accreted phosphides first reacted with liquid water, but then might have been cycled through hydrothermal systems where they would have been converted to more stable phosphates.

The time scale of our experiments is much shorter than those of water–rock reactions in Enceladus, giving rise to uncertainties in the exact phosphate/carbonate minerology inside the icy moon. Recent work[17] estimated $[\Sigma PO_4^{3-}]$ in Enceladus's ocean at low temperatures, as achieved by interactions involving thermodynamically stable phosphate and carbonate minerals (for example, merrillite and dolomite). The upper end of the predicted range from that study also agrees with the ocean concentration of phosphates (0.8–21 mM) derived here if the ocean pH is 8.0–10.5. This is consistent with the range of 9.0–10.5 adopted in this work. A notable difference is that the range given by ref. [17] also permitted $[\Sigma PO_4^{3-}]$ to reach much lower values—specifically at pH above 10.0—than those based on the hydroxyapatite-calcite buffer in Fig. 4a. The probable reason for this is that, under the strongly carbonated conditions modelled by ref. [17], hydroxyapatite was more soluble and thus merrillite took over the role of the P-controlling phosphate mineral. It is important to note that both systems produce results in good agreement with the CDA measurements and, in both cases, the driver enabling the abundant availability of phosphate is the high observed concentration of dissolved carbonate species, which shift phosphate-carbonate mineral equilibria toward dissolution of solid phosphates into Enceladus's ocean.

An earlier study by Lingam and Loeb[15] suggested that the availability of phosphorus would be the bottleneck of bio-essential elements on Enceladus and other icy ocean worlds with potential hydrothermal activity and without dry land. Steady-state concentrations “likely lower than the corresponding value on Earth by a few orders of magnitude”[15] would markedly reduce the prospects for life. Indeed, of the six elements—C, H, N, O, P and S—that are generally considered to be critical ingredients for life based on water and organic chemistry[12], phosphorus is cosmochemically the least abundant and has not previously been detected at any of the ocean-bearing moons in the Solar System. However, the results presented here demonstrate that Enceladus instead has a high availability of dissolved P, which is thus extremely unlikely to be a limiting factor in the survival of putative life on Enceladus—and perhaps also on other ocean worlds that reside beyond the CO_2 snowline in the Solar System (that is, the distance from the Sun beyond which CO_2 is in a solid (icy) state and is available as a planetary building material).

These wider implications are based on the fact that interactions between chondritic rocks and CO_2-rich fluids at a water:rock (W:R) ratio of around 1.0 would lead to a pH of about 10.0 (refs. [16,36]), in which dissolved phosphate concentrations tend to maximize[17] (Fig. 5). This is supported by recent findings for asteroid Ryugu, whose parent body is thought to have formed beyond the CO_2 snowline[37]. Given the high solubility of Na-bearing phosphate[17], its common occurrence in Ryugu samples is qualitatively consistent with our conclusion of phosphate-rich waters beyond the CO_2 snowline. Enceladus’s ocean could be a harbinger of high phosphorus availability in subsurface oceans across most of the outer Solar System. On the other hand, if an icy body formed within the CO_2 snowline both low [ΣCO_2] and high pH (approximately 13.0 at W:R around 1.0)[16,36] would probably limit dissolved phosphate concentrations (Fig. 5).

Fig. 5: Sensitivity of ΣPO_4^{3-} of the hydroxyapatite-calcite buffer system to pH and ΣCO_2 at 0.1 °C and 8 MPa.



The case of $\Sigma CO_2 = 0.2$ M (red curve) corresponds to the ocean water of Enceladus[3] and, perhaps, to icy bodies beyond the CO_2 snowline that originally contained CO_2 ice. The typical pH of such ΣCO_2-rich seawater is around 9.0–11.0 at W:R of approximately 1.0, owing to high concentrations of dissolved carbonates[16,36]. The case of $\Sigma CO_2 = 0.01$ M (purple curve) might correspond to ocean water of icy bodies within the CO_2 snowline that originally contained no CO_2 ice. According to ref. [16], ΣCO_2 in fluid becomes in the order of 0.01 M by interactions between CI chondritic rocks and pure water at 25 °C, around 1 bar and W:R approximately 1.0. The typical pH of such CO_2-poor water would be around 11.0–13.0 (refs. [16,36]). Red and blue shaded areas denote probable pH ranges for waters of icy bodies beyond and within the CO_2 snowline, respectively[16,36].

The stark contrast between earlier modelling and our results might be due to modelling assumptions that were based on scaling fluxes of the P cycle on the modern Earth to Enceladus[15], not considering the fundamental differences between Earth and ocean-bearing moons. The most important differences are the much higher concentration of carbonate species in alkaline ocean water and the probable presence of unrecycled, equilibrated rocks at the seafloor of Enceladus[34,38] versus continuous production of more reactive seafloor basalts on Earth. Regardless of these theoretical considerations, with the finding of phosphates the ocean of Enceladus is now known to satisfy what is generally considered to be the strictest requirement of habitability.

Methods

CDA data analysis

Short description of the CDA chemical analyser subsystem

The chemical analyser is the subsystem of the CDA[10] that provides compositional information about an impacting dust particle. Depending on the trajectory of the particle, it either hits the central rhodium target (chemical analyser target (CAT), diameter 0.17 m), the surrounding gold target (diameter 0.41 m) or the inner wall of the instrument. This work deals only with impacts on the CAT. If a dust particle impacts the CAT with sufficient energy, it is totally vaporized and partly ionized, forming an impact plasma of target and particle ions together with electrons, neutral molecules and atoms. The instrument separates this plasma, the positive component of which is linearly accelerated towards a multiplier about 19 cm away, used to generate a ToF spectrum. The spectrometer is sensitive to positive ions only. The mass resolution ($m/\Delta m$), derived from laboratory experiments with the instrument, depends on the atomic masses of the ions. At 1 u, $m/\Delta m$ is 10, increasing to 30 $m/\Delta m$ at 100 u and up to 50 $m/\Delta m$ at 190 u, although these values vary markedly with impact conditions.

Acquisition of a spectrum can be triggered by charge thresholds being exceeded on either the target or the multiplier ion detector (when a certain

abundant ion species arrives there, typically H^+, H_3O^+ or Na^+). In the case of a multiplier-trigger, the spectra show only ions with masses higher than the triggering ion species; the triggering species and those with lower masses do not appear in the spectrum. For this work, all P-rich spectra were triggered by impact. The spectra are logarithmically amplified, digitized at eight-bit resolution and sampled at 100 MHz for a period of 6.4 μs after the trigger. The recording period of the high-rate sampling mode allows the detection of ions with mass of up to 185–200 u, assuming that instrument recording is triggered by the impact itself.

Because ToF is proportional to the square root of the mass:charge ratio of ions, its spectrum in an ideal case also represents a mass spectrum for identical ion charges. The ions created by impact ionization in the impact speed regime considered here (roughly 10 km s^{-1}) are almost exclusively singly charged. Unfortunately, ToF is also influenced by the broad distribution of initial ion velocities, slightly varying flight paths and plasma-shielding effects[40]. For that reason, species of identical mass are distributed over a range of sampling points and the mass resolution drops below integer values, usually at 20–30 u and above. Therefore, we prefer to show original CDA spectra with an *x* axis showing ToF rather than mass; the latter could mislead the reader to intuitively assume an unrealistically high accuracy.

For a detailed description of the instrument see ref. [10]; the calibration routine is described in refs. [6,20]. For the dataset in this work the mass calibration was generally done using sodium (Na^+) and sodium hydroxide ($(NaOH)Na^+$) mass lines as reference and a stretch factor of $a = 473$ ns.

CDA dataset

All data used for this work are listed in Extended Data Table 1 and, like all CDA data, are archived on the Small Bodies Node of the Planetary Data System (PDS–SBN), at https://sbn.psi.edu/pds/resource/cocda.html. We used 15 time periods between 2004 and 2008 covering times of CDA pointing that were favourable for the detection of E-ring dust. Times this early in the Cassini tour were chosen because CDA contamination from salts deposited on the CDA impact target during deep Enceladus plume crossings was then negligible[39]. Within these periods a total of 7,353 spectra from E-ring grains

were recorded. A Lee filter was applied to the spectra to improve signal to noise. Of these spectra, 962 have been categorized as Type 3 (Extended Data Table 1).

From this set of Type 3 spectra, nine stand out as belonging to a previously unknown subtype with a unique pattern of three peaks of *m*/*z* above 120 u; these are the P-rich spectra discussed above. Individual events are listed in Extended Data Table 2. Remarkably, all these spectra were triggered on the instrument by the impact charge of the particle whereas only 36% of all Type 3 spectra in the dataset were triggered in this way (Extended Data Table 1). The remaining 64% of Type 3 spectra were triggered by a particular ion species reaching the multiplier (Methods). By applying binomial statistics, we get a probability that all P-rich spectra were impact charge triggered by chance of only 0.01%. Therefore, we assume a systematic effect in spectral recognition for P-rich grains. Indeed, it is plausible that the spectrum of a P-rich grain is not identified as Type 3 if it is triggered by an ion species at the multiplier. Type 3 spectra are in most cases triggered by Na^+, which is then missing in the spectrum. However, without the Na^+ line and in the absence of other typical Type 3 features such as $(NaCl)_nNa^+$ or $(Na_2CO_3)Na^+$, P-rich spectra are not identifiable as being Type 3 following the criteria normally used, which focus on spectra dominated by chloride, carbonate and hydroxide peaks[3,11].

We therefore used only those 345 spectra of the dataset that—like the new phosphate-rich spectra—have been triggered by impact (Extended Data Table 1) as a reference to calculate the frequency of the new phosphate-bearing subtype within the family of Type 3 grains and, with that, the average phosphate concentration in salt-rich Type 3 grains. It is noteworthy though, that even in the unlikely case that there was no such systematic selection effect and we use the entire dataset for the inference of average phosphate concentration, the lower limit will still be ΣPO_4^{3-} above 0.3 mM, more than 100-fold the mean concentration in Earth's oceans, and the conclusions of this work are unaffected.

Laboratory methods simulating CDA mass spectra

Identification of sodium phosphates

The largest peaks in all nine CDA spectra always appear at 23 and 63 u, indicative of Na^+ and $(NaOH)Na^+$ cations that are the defining features in all salt-rich spectra from grains emitted by Enceladus (Type 3; ref. [3]), thus these spectra contain sodium salts in an alkaline matrix. We used the other three characteristic peaks, at 125, 165 and 187 u (allowing a divergence of ±1 u that reflects the mass uncertainty of CDA[7,10]), to perform a rigorous database search—in both our own database base[41] as well as several external databases (for example, NIST Chemistry WebBook, MassBank North America and MassBank North Europe (https://massbank.eu/MassBank)). We found phosphates to be the only possible class of Na-rich compounds that produce a pattern that is in even remote agreement with the observed cationic peak pattern. After that we widened the accepted mass range to ±2 u away from the observed masses and relaxed the restriction to Na-rich compounds to determine whether there are other species (without sodium) forming a characteristic mass line pattern similar to the three mass lines (125, 165 and 187 u). The only match remotely close was methyl-2,2-dimethyl-3-(2,2-dichloroethenyl) cyclopropanoate, showing prominent peaks at m/z 61(−2), 127(+2), 163(−2) and 187. The mass spectrum can be found at ref. [42].

It is questionable whether a sufficient concentration of this type of chlorinated organic ester would be stable in Enceladus's alkaline ocean. Even if we assume that this unlikely substance could be mixed with sodium salts in an ice grain, the spectrum has additional features not matching the CDA spectrum—most prominently a peak at m/z 91, which is absent in all nine CDA spectra. Similarly, any combination of exotic organic compounds we found that remotely reproduces the five prominent CDA peaks does have a number of additional peaks that are absent in CDA spectra.

In conclusion, sodium phosphates are the simplest and—to the best of our knowledge—the only explanation for all the peaks observed in the CDA spectra (Figs. 1 and 2). These salts have already been predicted to be relatively abundant in Enceladus's ocean based on recent geochemical modelling[17], and the observed peak pattern can be readily reproduced in mass spectra from our analogue experiment, as described in the following section.

LILBID experiments simulating CDA spectra

We used the laser-induced liquid beam ion desorption (LILBID) facility at Freie Universität Berlin to simulate CDA spectra. The experimental setup is described in detail in ref. [22] and here we provide only a brief overview. A micrometre-sized water beam is irradiated by a pulsed infrared laser (20 Hz, 7 ns pulse length) at a wavelength of 2,840 nm and variable laser energy. On absorbing the laser energy, the water beam explosively disperses into neutral and charged macroscopic, molecular and elemental fragments. Cations or anions, depending on the instrument polarity, are accelerated and analysed in a ToF mass spectrometer. Detected signals are amplified, digitized and recorded with a LabVIEW-controlled computer. Typically 300–500 individual spectra are co-added and averaged during one effective measurement. The laboratory setup is calibrated before every measurement using a 10^{-6} M NaCl solution[22]. The recorded mass spectra showed a mass resolution of 600–800 $m/\Delta m$ (full-width at half-maximum).

As discussed in Methods, we found phosphates to be the only Na-rich compounds that are in even remote agreement with the observed cationic peak pattern. Because the appearance of certain peaks and their amplitudes are very sensitive to the phosphate species used and their concentrations, we individually tested a variety of combinations of aqueous solutions of Na_3PO_4, Na_2HPO_4, NaH_2PO_4 and Na_2HPO_3, as well as mixtures of these compounds at different concentrations with the goal of reproducing the CDA peak pattern. Concentrations were varied between 0.01 M and about 5 M for NaH_2PO_4, 0.5 M for Na_2HPO_4 and 1.5 M for Na_3PO_4.

We found that a mixture of Na_3PO_4 and Na_2HPO_4 most closely reproduced the CDA peak pattern of P-rich spectra (Fig. 1), whereas spectra of the pure substances always exhibited differences (Extended Data Figs. 1 and 2).

We also tested for the presence of phosphites (using Na_2HPO_3, in which P has an oxidation state of +III) and hypophosphite (using NaH_2PO_2, in which P has an oxidation state of +I). Both substances created mass lines not in agreement with CDA observations: $(Na_2HPO_3)Na^+$ (149 u) for phosphite and $(NaH_2PO_2)Na^+$ (111 u) for hypophosphite. For phosphite we inferred an

upper limit concentration of 5 mM to be in agreement with the absence of a mass line at 149 u in CDA spectra (Extended Data Fig. 3).

The phosphate salts used in our experiments were:

- Na_3PO_4: Sigma-Aldrich, 96% purity
- Na_2HPO_4: Roth, over 99% purity
- $NaH_2PO_4{\bullet}2H_2O$: AppliChem, over 98% purity
- $NaH_2PO_3{\bullet}5H_2O$: Sigma-Aldrich, over 98% purity
- $NaH_2PO_2{\bullet}H_2O$: Sigma-Aldrich, over 99% purity

Theoretical considerations regarding the observed characteristic pattern of phosphate ions

We observed the peak sequence at masses 187, 165 and 125 u as a characteristic pattern indicating the presence of phosphorus (P) salts in Enceladus's ice grains, present in all nine of the individual CDA ice grain spectra. Phosphate-containing salts from the solvation and extraction of minerals at a basic pH value produced a number of anion species in solution that are coupled through a multistep acid–base thermodynamic equilibrium and proton exchange—for example, $PO_4^{3-} \leftrightarrow HPO_4^{2-} \leftrightarrow H_2PO_4^{-}$. For the given high salt concentrations, abundance of counter cations are close to anion abundances in solution ($Na_nPO_4^{(3-n)-}$). After ice grain impact, ionic aggregates showed additional clustering with abundant cations such as Na^+ resulting in characteristic mass peaks due to species including $(Na_2HPO_4)Na^+$ and $(Na_3PO_4)Na^+$. Owing to the basic pH of the solution forming the ice grains (or the liquid desorption matrix of the laboratory experiment), these two species are expected.

However, we did not detect a $(NaH_2PO_4)Na^+$ adduct in CDA spectra but rather a $(NaPO_3)Na^+$ peak at 125 u, a charged molecular aggregate of a metaphosphate with phosphorus in the same oxidation state (+V) as the

other orthophosphates. We attribute the absence of $(NaH_2PO_4)Na^+$ and the appearance of $(NaPO_3)Na^+$ instead to an elimination reaction during desorption (LILBID) or particle impact (CDA):

NaH_2PO_4 + heat/energy → $NaPO_3 + H_2O$, followed by the addition of Na^+.

The fact that the $(NaPO_3)Na^+$ peak at 125 u is formed from hydrogen phosphates (linked via an acid–base equilibrium) in our analogue experiments, but not from Na_3PO_4 alone, supports this picture. However, it should be noted that we could produce a $(NaH_2PO_4)Na^+$ peak (143 u) if high concentrations (over around 0.05 M) of NaH_2PO_4 or Na_2HPO_4 (Extended Data Fig. 2) were used in the laboratory experiments or if the pH value of the solution was forced below about 9.0. We then find a clear correlation of the amplitudes of dihydrogen phosphate (143 u) with metaphosphate (125 u). In the CDA data the 143 u peak appears to be below the noise level of the detector. In our laboratory spectra the 125 u metaphosphate peak is generally three- to fivefold more intense and it is therefore likely that small amounts of $(NaH_2PO_4)Na^+$ (143 u) were buried in the noise of the CDA spectrum. Such a signature was indeed tentatively observed in one of the nine spectra (event 3 in Extended Data Table 1).

To test these ideas we calculated the thermochemical properties of the water elimination of NaH_2PO_4 + heat/energy → $NaPO_3 + H_2O$ with the Jaguar quantum chemistry programme package at the PBE0-D3/6-311+G(d,p)/PBE level of theory. Two variants have been calculated: (1) with Na^+ as a counter ion and (2) without Na^+ (involving a $H_2PO_4^-$ anion). The reaction is slightly endergonic in the gas phase and even more favoured in/with water. The barrier of the concerted elimination reaction appears to be low if water molecules are involved. One may even consider it to be a water-assisted reaction, in particular for the case of Na^+ as counter ion (Extended Data Fig. 4).

Hydrothermal experiments and calculations of chemical equilibria

Experimental methods and procedures

Hydrothermal experiments were performed with a Dickson-type hydrothermal apparatus (Toyo Koatsu Co. Ltd) used in a previous study[43]. In this apparatus, water–rock reactions occur in a flexible reaction cell that comprises a gold bag with a titanium head. The flexible reaction cell was connected to a stainless steel sampling tube with a gold-coated inner wall, to avoid catalytic reactions. The reaction cell was heated at 500 °C for 3 h in air before each experiment to remove potential contamination by organic matter. The reaction cell was placed in a stainless steel autoclave, which was held at a pressure and temperature of 30 MPa and 150 °C, respectively, during the experiments. The reaction cell was pressurized by the addition of surrounding water into the autoclave using a hand pump. The autoclave was heated with an external heater. Due to the flexibility of the gold bag, the reaction cell deforms plastically without fracturing in response to the addition of surrounding water. This allowed us to undertake online sampling of fluid without significant changes in pressure.

The starting rock material was prepared by powdering a few pieces of the commercially available Jbilet Winselwan carbonaceous chondrite (CM-2 type)[25] in an agate mortar. The starting solutions were prepared by dissolving $NaHCO_3$ powder and NH_3 solution in ultrapure water such that the initial ΣCO_2 was 0.2 M for Run no. 1 and 1.0 M for Run no. 2, and ΣNH_3 was 0.5 M. One way to achieve high ΣCO_2 in an Enceladus experiment is to introduce gaseous CO_2 in the reaction cell and dissolve it by pressurization. However, this is technically challenging in our experimental system and we cannot control the $[\Sigma CO_2]$ and pH of the starting fluids. Therefore, we achieved high ΣCO_2 by the addition of $NaHCO_3$ powder. We added 1.0 M Na in Run no. 2, higher than Na concentrations in Enceladus's ocean[3]. If Na phosphate (for example, Na struvite) had precipitated in such Na-rich fluids of Run no. 2, this would have affected the buffer system that controls ΣPO_4 in fluids. Although the temperature and pressure dependences of the thermodynamic constants of Na struvite have not been experimentally determined, the saturation index ($\log([Na^+][Mg^{2+}][HPO_4^{2-}]/K_{\text{Na-struvite}})$, where $K_{\text{Na-struvite}}$ denotes the

equilibrium constant from ref. [17] of Na struvite at 1 bar and 25 °C, is calculated to be below −1.6 using the measured Na, Mg and ΣPO_4 concentrations and in situ pH. Because the fluid Mg concentrations of Run no. 2 were below the detection limit (Extended Data Fig. 8b), Na struvite was likely to have been undersaturated in the experiment. The presence of Ca sulfate in the starting rock materials (CM chondrite) also does not affect our conclusion regarding the hydroxyapatite-calcite buffer system. If the hydroxyapatite-gypsum (or hydroxyapatite-anhydrite) buffer controls ΣPO_4^{3-}, [ΣPO_4^{3-}] should dramatically increase with [SO_4^{2-}]; however, this did not occur (Extended Data Table 3). The initial amounts of starting rock powder and solution were around 7 g and 10 ml, respectively. In each experiment we collected several aliquots (about 1.2 ml) of fluid for analysis. As such, water:rock mass ratios in the reaction cell changed from around 1.4–0.9 due to the removal of multiple fluid samples.

The collected fluid samples were analysed by inductively coupled plasma atomic emission spectroscopy (ICP–AES: SPS5510, Hitachi High-Tech) for dissolved Na, K, Mg, Ca, Si and Al content; by ICP mass spectrometry (ICP–MS: Agilent 8000) for dissolved Fe content; by ion chromatography (ICS-1600, DIONEX) for dissolved Cl and SO_4 content; and by gas chromatography (GC-2010, Shimadzu) for ΣCO_2 content (Extended Data Table 3). ΣPO_4^{3-} content was analysed by both tandem ICP–MS (ICP–MS/MS: Agilent 8000, Agilent) and ion chromatography. The former methodology quantified dissolved total P in fluids at high sensitivity, including phosphate and potentially phosphite, whereas the latter directly measured ΣPO_4^{3-} with moderate sensitivity. The results of the two analyses showed that, basically, all dissolved total P in the fluids was phosphate (Extended Data Table 3). As such, we used ICP–MS/MS results to quantify ΣPO_4^{3-} in the fluid samples. In situ pH values at 150 °C during the experiments were calculated using the geochemical code PHREEQC v.3 (ref. [44]), based on the measured pH at 25 °C and dissolved elements, molecules (phosphate and sulfate) and ΣCO_2 in the diluted fluid samples. To calculate the in situ pH of the alkaline, Na-carbonate-rich fluids, we constrained the charge balance from the measured pH at 25 °C in two ways. First, Na was used to compensate for charge imbalance due to analytical

uncertainties. Second, ΣCO_2 was used to compensate for charge imbalance. We adopted the median value between the two values as the in situ pH, and the difference between the median and upper/lower values as the error. We also conducted mineralogical and chemical analyses of solid residues after the experiments using an X-ray diffraction spectrometer (XRD: MiniFlex600, Rigaku) and a scanning electron microscope with an electron probe microanalyser (EPMA: JXA-8530F, JEOL).

The ΣPO_4^{3-} concentrations controlled by the hydroxyapatite-calcite and whitlockite-calcite buffer systems were calculated with the equilibrium constants obtained from the SUPCRT92 programme[45]. In the calculations, the thermodynamic properties of whitlockite (β-$Ca_3(PO_4)_2$) and hydroxyapatite were taken from Robie et al.[46].

Elemental composition of carbonate and phosphate

Extended Data Figs. 5 and 6 show typical secondary electron microprobe images and elemental mapping for thick sections of solid residues after the experiments (Run nos. 1 and 2). In the solid residues we found large grains of precipitated Ca carbonate ($CaCO_3$, size over 100 μm; Extended Data Fig. 5 and Extended Data Table 4). Because the typical size of Ca carbonate in the sample before the experiments was 10 μm or less (Extended Data Fig. 7), the large Ca carbonate grains found after the experiments are highly likely to have been precipitated during the experiments through ΣCO_2 sequestration. This view is supported by decreases in ΣCO_2 in fluids (Extended Data Fig. 8). Our XRD analysis shows that the precipitated Ca carbonates are most likely calcite in both of the experiments (Run nos. 1 and 2; Extended Data Fig. 10).

To investigate the composition of the phyllosilicate phase before and after the experiments, we performed EPMA analysis for the phyllosilicate matrix around Ca phosphate grains. Extended Data Fig. 9 shows a ternary diagram with the results of EPMA analysis for phyllosilicate, Ca phosphate and Fe oxides in the samples before and after the experiments. Phyllosilicates in the starting materials were mostly serpentine; however, mixtures of serpentine and saponite appeared in the samples after the experiments. Extended Data

Fig. 9 indicates a linear trend of a mixing line of the compositions between Ca phosphate and serpentine (and saponite mixtures).

Extended Data Table 4 presents the results of spot analysis for Ca phosphate and carbonate found in the solid residues after the experiments of Run nos. 1 and 2. At the micrometre scale, Ca phosphate minerals usually coexist with other secondary minerals such as serpentine and saponite (Extended Data Fig. 6), and thus the results of the spot analysis might not be explained by the simple chemical composition of an alteration mineral. As reported in Extended Data Table 4, Grain C with the highest P content had high total mass fractions of around 95–100 wt%, suggesting that the contribution of calcite to Ca:P atomic ratios is small (less than 2–3% of Ca mole abundance), whereas serpentine and saponite would coexist with Ca phosphate of Grain A, Vein B and Grain D. These hydrated secondary minerals can reduce the total mass fractions of spot analyses by 5–10% owing to the presence of H_2O and OH. Because the total mass fractions for Grain A, Vein B and Grain D are also high in general (over 87 wt%; Extended Data Table 4), calcite is presumed to be rare in Ca phosphate grains, suggesting that the contribution of calcite to Ca:P ratios is small.

Grain C of Run no. 2 has very high P:Si ratios (Extended Data Table 4c). The measured Ca:P mole ratios are 1.32–1.37. Grain D of Run no. 2 has relatively high P:Si ratios, above 1.0 (Extended Data Table 4d). The Ca:P mole ratios of Grain D range from 1.29 to 1.34. The Ca:P ratios of hydroxyapatite, whitlockite [β-$Ca_3(PO_4)_2$] and Mg-whitlockite [$Ca_9MgH(PO_4)_7$] are 1.67, 1.50 and 1.29, respectively. Ca phosphates collected after Run no. 2 (Grains C and D) are inferred to consist of mixtures of multiple Ca phosphates. Due to the trace amounts of Ca phosphates in the solid residues, these could not be identified in the XRD spectra (Extended Data Fig. 10)

In Ca phosphate grains of Run no. 1 (Grain A and Vein B), P:Si mole ratios are generally low (roughly 0.1–1.0) (Extended Data Table 4a,b), suggesting that Ca phosphates coexist with serpentine and saponite. There is a relatively wide variation in the Ca:P mole ratios of Grain A and Vein B, from approximately 1.1 to 1.5 (Extended Data Table 4a,b). A high Ca:P ratio of around 1.5 requires the presence of hydroxyapatite and/or whitlockite. To

explain the low Ca:P ratios (below 1.29) of the analysed spots on Grain A and Vein B, the presence of Ca-poor phosphate, such as NH_4MgPO_4, would be needed.

Our results showing high C:P ratios (1.32–1.37) for Run no. 2 and low Ca:P ratios (1.25–1.27) for Run no. 1 are consistent with our interpretation of the buffer system that controls ΣPO_4^{3-} concentrations in the fluid samples. The Ca phosphate grains (C and D) of Run no. 2 contain high fractions of phosphate with high Ca:P, such as hydroxyapatite. This observation is consistent with the control of ΣPO_4^{3-} concentrations in fluid samples of Run no. 2 by the hydroxyapatite-calcite buffer (Fig. 3). In contrast, the observation of Ca phosphate grains with low Ca:P for Run no. 1 (for fluid with ΣCO_2 = 0.01 M and reaction time 1,000 h) is consistent with the control of ΣPO_4^{3-} concentrations by Ca phosphate with lower Ca:P ratios, such as whitlockite, rather than control by hydroxyapatite (Fig. 3). This fluid sample was collected in the final sampling of Run no. 1 (Extended Data Fig. 8a), and the solid residue was collected after this final sampling. We consider that Ca phosphate with relatively low Ca:P, such as whitlockite [β-$Ca_3(PO_4)_2$], Mg-whitlockite and $NaMgPO_4$, would have been precipitated before the final sampling in response to an increase in Mg concentration in the fluid (Extended Data Fig. 8), and that the whitlockite-calcite buffer would have controlled the P concentration of the fluid at the end of experimental Run no. 1.

Chemical equilibrium considerations regarding phosphate buffering systems

Extended Data Fig. 11 shows the calculated results of [ΣPO_4^{3-}] in fluids as a function of temperature for different pH and [ΣCO_2] values, assuming the whitlockite-calcite buffer system (see above for the equation). Although our experimental data suggest that the whitlockite-calcite buffer could have controlled ΣPO_4^{3-} concentrations in the fluid at the end of Run no. 1, we consider that this could have been an experimental artefact reflecting transient states of supersaturation of Mg-bearing phases. In the experiments, the dissolution of anhydrous primordial minerals (for example, olivine and

pyroxene) would have provided Mg and Ca for the fluids. These Mg and Ca ions would have been removed by reactions with CO_3^{2-}, forming carbonate minerals (Extended Data Fig. 5). However, at the end of Run no. 1, ΣCO_2 in the fluids was depleted (Extended Data Fig. 8) owing to CO_2 sequestration via carbonate mineral formation; therefore, dissolved Mg and Ca concentrations increased in the fluids (Extended Data Fig. 8). In longer-period reactions, excess Mg and Ca ions would have been consumed through the formation of serpentine and saponite. Thermodynamic equilibrium calculations show that Mg and Ca concentrations in fluids that interact with carbonaceous chondritic rocks are typically low (0.001–0.010 mM each) in alkaline fluids (pH over 9.0) in the presence of serpentine and saponite[36,47]. In Enceladus, water–rock reactions would be driven closer to acid–base equilibrium because the reactions can occur over much longer time scales than apply in our experiments. If the ocean of Enceladus were consistent with the expectation of low Mg and Ca concentrations (for example, under 0.1 mM each), then ΣPO_4^{3-} in fluids in contact with chondritic rocks would be controlled by the hydroxyapatite-calcite buffer system.

In this study, we considered hydroxyapatite and whitlockite [β-$Ca_3(PO_4)_2$] as phosphate minerals of buffer systems that control [ΣPO_4^{3-}] in fluids. However, previous work has suggested the possibility of the occurrence of other phosphate minerals in water–rock reactions in Enceladus[17], including Mg-whitlockite and merrillite. Using the equilibrium constants of dissolution of Mg-whitlockite and merrillite at 1 bar and 25 °C (ref. [48]), and assuming [Mg^{2+}] as roughly 10^{-5} M in the fluid samples (Extended Data Table 3), ΣPO_4^{3-} values controlled by the Mg-whitlockite-calcite and merrillite-calcite buffer systems (that is, 6×10^{-5} and 7×10^{-6} M, respectively, for $\Sigma CO_2 = 0.1$ M) are about 1/30 to 1/250 that of the hydroxyapatite-calcite buffer (that is, 2×10^{-3} M for $\Sigma CO_2 = 0.1$ M). Although the temperature and pressure dependences of the thermodynamic equilibrium constants of both Mg-whitlockite and merrillite are unknown, the observed ΣPO_4^{3-} values in our experiments are unlikely to be explained by these buffer systems.

We also assumed that calcite is the carbonate mineral phase for the buffer system, but dolomite is usually more abundant than calcite in CI chondrites[49]. However, ΣPO_4^{3-} values controlled by the hydroxyapatite-dolomite buffer are orders of magnitude higher than those controlled by the hydroxyapatite-calcite buffer for $\Sigma CO_2 = 0.1$ M and the typical Mg concentration in our experiments ($[Mg^{2+}] = 10\ M^{-5}$). Therefore, we conclude also that the hydroxyapatite-dolomite buffer may not explain the observed ΣPO_4^{3-} in our experiments.

Data availability

All CDA data used for this work are listed in Extended Data Table 1 and are archived on PDS–SBN, at https://sbn.psi.edu/pds/resource/cocda.html. Data from the LILBID analogue experiment reproducing CDA data and geochemical experiments are available in the Zenodo public repository at https://doi.org/10.5281/zenodo.7703848.

References

1. Thomas, P. C. et al. Enceladus's measured physical libration requires a global subsurface ocean. *Icarus* **264**, 37–47 (2016).

2. Iess, L. et al. The gravity field and interior structure of Enceladus. *Science* **344**, 78–80 (2014).

3. Postberg, F. et al. Sodium salts in E-ring ice grains from an ocean below the surface of Enceladus. *Nature* **459**, 1098–1101 (2009).

4. Hsu, H.-W. et al. Ongoing hydrothermal activities within Enceladus. *Nature* **519**, 207–210 (2015).

5. Waite, J. H. et al. Cassini finds molecular hydrogen in the Enceladus plume: evidence for hydrothermal processes. *Science* **356**, 155–159 (2017).

6. Postberg, F. et al. The E-ring in the vicinity of Enceladus: II. Probing the moon's interior—the composition of E-ring particles. *Icarus* **193**, 438–454 (2008).

7. Postberg, F. et al. Macromolecular organic compounds from the depths of Enceladus. *Nature* **558**, 564–568 (2018a).

8. Postberg, F. et al. in *Enceladus and the Icy Moons of Saturn* (eds Schenk, P. M. et al.) 129–162 (Univ. Arizona Press, 2018).

9. Khawaja, N. et al. Low-mass nitrogen-, oxygen-bearing, and aromatic compounds in Enceladean ice grains. *Mon. Not. R. Astron. Soc.* **489**, 5231–5243 (2019).

10. Srama, R. et al. The Cassini Cosmic Dust Analyzer. *Space Sci. Rev.* **114**, 465–518 (2004).

11. Postberg, F. et al. A salt-water reservoir as the source of a compositionally stratified plume on Enceladus. *Nature* **474**, 620–622 (2011).

12. Cockell, C. S. et al. Habitability: a review. *Astrobiology* **16**, 89–117 (2016).

13. Westheimer, F. H. Why nature chose phosphates. *Science* **235**, 1173–1178 (1987).

14. Tyrell, T. The relative influences of nitrogen and phosphorus on oceanic primary production. *Nature* **400**, 525–531 (1999).

15. Lingam, M. & Loeb, A. Is extraterrestrial life suppressed on subsurface ocean worlds due to the paucity of bioessential elements? *Astron. J* **156**, 151 (2018).

16. Zolotov, M. Y. Aqueous fluid composition in CI chondritic materials: chemical equilibrium assessments in closed systems. *Icarus* **220**, 713–729 (2012).

17. Hao, J. et al. Abundant phosphorus expected for possible life in Enceladus's ocean. *Proc. Natl Acad. Sci. USA* **119**, e2201388119 (2022).

18. Choblet, G. et al. Powering prolonged hydrothermal activity inside Enceladus. *Nat. Astron.* **1**, 841–847 (2017).

19. Liao, Y., Nimmo, F. & Neufeld, J. A. Heat production and tidally driven fluid flow in the permeable core of Enceladus. *J. Geophys. Res. Planets* **125**, e2019JE006209 (2020).

20. Postberg, F. et al. Composition of jovian dust stream particles. *Icarus* **183**, 122–134 (2006).

21. Kempf, S. et al. in *Enceladus and the Icy Moons of Saturn* (eds Schenk, P. M. et al.) 195–2018 (Univ. Arizona Press, 2018).

22. Klenner, F. et al. Analogue spectra for impact ionization mass spectra of water ice grains obtained at different impact speeds in space. *Rapid Commun. Mass Spectrom.* **33**, 1751–1760 (2019).

23. Wiederschein, F. et al. Charge separation and isolation in strong water droplet impacts. *Phys. Chem. Chem. Phys.* **17**, 6858–6864 (2015).

24. Shock, E. L., Sassani, D. C., Willis, M. & Sverjensky, D. A. Inorganic species in geologic fluids: correlations among standard molal thermodynamic properties of aqueous ions and hydroxide complexes. *Geochim. Cosmochim. Acta* **61**, 907–950 (1997).

25. King, A. J. et al. The alteration history of the Jbilet Winselwan CM carbonaceous chondrite: an analog for C-type asteroid sample return. *Meteorit. Planet. Sci.* **54**, 521–543 (2019).

26. Sekine, Y. et al. High-temperature water–rock interactions and hydrothermal environments in the chondrite-like core of Enceladus. *Nat. Commun.* **6**, 8604 (2015).

27. Glein, C. R., Baross, J. A. & Waite, J. H. The pH of Enceladus' ocean. *Geochim. Cosmochim. Acta* **162**, 202–219 (2015).

28. Fukushi, K. et al. In situ formation of monohydrocalcite in alkaline saline lakes of the Valley of Gobi Lakes: Prediction for Mg, Ca, and total dissolved carbonate concentrations in Enceladus' ocean and alkaline-carbonate ocean worlds. *Minerals* **10**, 669 (2020).

29. Toner, J. D. & Catling, D. C. A carbonate-rich lake solution to the phosphate problem of the origin of life. *Proc. Natl Acad. Sci. USA* **117**, 883–888 (2020).

30. Mautner, M. N. & Sinaj, S. Water-extractable and exchangeable phosphate in Martian and carbonaceous chondrite meteorites and in planetary soil analogs. *Geochim. Cosmochim. Acta* **66**, 3161–3174 (2002).

31. Jang, H. et al. Revisiting whitlockite, the second most abundant biomineral in bone: nanocrystal synthesis in physiologically relevant conditions and biocompatibility evaluation. *ACS Nano* **8**, 634–641 (2014).

32. Berner, E. K. and Berner, R. A. *Global Environment: Water, Air, and Geochemical Cycles* (Princeton Univ. Press, 2012).

33. Pasek, M. A. & Lauretta, D. S. Aqueous corrosion of phosphide minerals from iron meteorites: a highly reactive source of prebiotic phosphorus on the surface of the early Earth. *Astrobiology* **5**, 515–535 (2005).

34. Glein, C. R., Postberg, F. & Vance, S. in *Enceladus and the Icy Moons of Saturn* (eds Schenk, P. M. et al.) 39–56 (Univ. Arizona Press, 2018).

35. Daval, D., Choblet, G., Sotin, C. & Guyot, F. Theoretical considerations on the characteristic timescales of hydrogen generation by serpentinization reactions on Enceladus. *J. Geophys. Res. Planets* **127**, e2021JE006995 (2022).

36. Kurokawa, H. et al. Distant formation and differentiation of outer main belt asteroids and carbonaceous chondrite parent bodies. *AGU Adv.* **3**, e2021AV000568 (2022).

37. Nakamura, T. et al. Formation and evolution of carbonaceous asteroid Ryugu: direct evidence from returned samples. *Science* **379**, eabn8671 (2022).

38. Glein, C. R. & Waite, J. H. The carbonate geochemistry of Enceladus' ocean. *Geophys. Res. Lett.* **47**, e2019GL085885 (2020).

39. Postberg, F. et al. Discriminating contamination from particle components in spectra of Cassini's dust detector CDA. *Planet. Space Sci.* **57**, 1359–1374 (2009b).

40. Hillier, J. K., McBride, N., Green, S. F., Kempf, S. & Srama, R. Modelling CDA mass spectra. *Planet. Space Sci.* **54**, 1007–1013 (2006).

41. Klenner, F. et al. Developing a laser induced liquid beam ion desorption spectral database as reference for spaceborne mass spectrometers. *Earth Space Sci.* **9**, e2022EA002313 (2022).

42. KURARAY CO., LTD. MassBank Record: MSBNK-Fac_Eng_Univ_Tokyo-JP009911. https://massbank.eu/MassBank/RecordDisplay?id=MSBNK-Fac_Eng_Univ_Tokyo-JP009911 (2011)

43. Tan, S., Sekine, Y., Shibuya, T., Miyamoto, C. & Takahashi, Y. The role of hydrothermal sulfate reduction in the sulfur cycles within Europa: laboratory experiments on sulfate reduction at 100 MPa. *Icarus* **357**, 114222 (2021).

44. Parkhurst, D. L. & Appelo, C. A. J. Description of input and examples for PHREEQC version 3—A computer program for speciation, batch-reaction, one-dimensional transport, and inverse geochemical calculations. In *U.S. Geological Survey Techniques and Methods,* book 6, Ch. A43, 497 (USGS, 2013).

45. Johnson, J. W., Oelkers, E. H. & Helgeson, H. C. SUPCRT92: a software package for calculating the standard molal thermodynamic

properties of minerals, gases, aqueous species, and reactions from 1 to 5000 bar and 0 to 1000 °C. *Comput. Geosci.* **18**, 899–947 (1992).

46. Robie, R. A., Hemingway, B. S. & Fisher, J. R. Thermodynamic properties of minerals and related substances at 298.15 K and 1 bar (105 pascals) pressure and at higher temperatures. *USGS Bull*. **1452**, 452 (1978).

47. Zolotov, M. Y. An oceanic composition on early and today's Enceladus. *Geophys. Res. Lett.* **34**, L23203 (2007).

48. Adcock, C. T., Hausrath, E. M. & Forster, P. M. Readily available phosphate from minerals in early aqueous environments on Mars. *Nat. Geosci.* **6**, 824–827 (2013).

49. Brearley, A. J. in *Meteorites and the Early Solar System II* (eds Lauretta, D. S. et al.) 584–624 (Univ. Arizona Press, 2006).

Acknowledgements

F.P., F.K., Z.Z., J.K.H., N.K. and L.N. were supported by the European Research Council under the European Union's Horizon 2020 research and innovation programme (Consolidator grant no. 724908-Habitat OASIS). Y.S. and T.Shibuya have been supported by Grants-in-Aid for Scientific Research from the Japan Society for Promotion of Science (grant nos. JP22K21344, JP17H06455 and JP17H06456). R.S. has been supported by the German Aerospace Centre through grant no. 50OH1501. B.A. acknowledges support by the German Science Foundation through grant no. AB 63/25-1. Z.Z. was supported by the Chinese Scholarship Council (grant no. 201604910894). C.R.G. was supported by Southwest Research Institute (grant no. 15-R6248) and by Simons Collaboration on the Origins of Life (grant no. 511570FY20).

Funding

Open access funding provided by Freie Universität Berlin.

Author information

Authors and Affiliations

1. Institut für Geologische Wissenschaften, Freie Universität Berlin, Berlin, Germany

   Frank Postberg, Fabian Klenner, Zenghui Zou, Jon K. Hillier, Nozair Khawaja, Lenz Noelle & Juergen Schmidt

2. Earth-Life Science Institute (ELSI), Tokyo Institute of Technology, Tokyo, Japan

   Yasuhito Sekine, Kento Furuya & Shuya Tan

3. Institute of Nature and Environmental Technology, Kanazawa University, Ishikawa, Japan

   Yasuhito Sekine

4. Space Science Division, Space Sector, Southwest Research Institute, San Antonio, TX, USA

   Christopher R. Glein

5. Leibniz-Institute für Oberflächenmodifizierung, Leipzig, Germany

   Bernd Abel

6. Institute of Chemical Technology, University of Leipzig, Leipzig, Germany

   Bernd Abel

7. Laboratory for Atmospheric and Space Physics (LASP), University of Colorado, Boulder, CO, USA

   Sascha Kempf

8. Institute for Extra-cutting-edge Science and Technology Avantgarde Research (X-star), Japan Agency for Marine-Earth Science and Technology (JAMSTEC), Kanagawa, Japan

   Takuya Saito & Takazo Shibuya

9. Astronomy Research Unit, University of Oulu, Oulu, Finland

   Juergen Schmidt

10. Institut für Raumfahrtsysteme, Universität Stuttgart, Stuttgart, Germany

   Ralf Srama

Contributions

F.P. led the study and writing of the manuscript, with support from Y.S., C.R.G., J.S. and J.K.H. Z.Z. and F.P. were responsible for CDA data analysis, with support from L.N. R.S. and S.K. designed the CDA observation and programmed codes for CDA data reduction. F.K. led the LILBID analogue experiments, with support from Z.Z. Y.S., K.F., T.Saito., T.Shibuya and S.T. conducted all geochemical experiments and carried out geochemical modelling, supported by C.R.G. B.A. was responsible for setting up and running of the quantum chemistry code. All authors contributed to the discussion and commented on the manuscript.

Corresponding author

Correspondence to Frank Postberg.

Ethics declarations

Competing interests

The authors declare no competing interests.

Peer review

Peer review information

Nature thanks Jeffrey Seewald, Michael Zolensky, Mikhail Zolotov and the other, anonymous, reviewer(s) for their contribution to the peer review of this work.

Additional information

Publisher’s note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Extended data figures and tables

Extended Data Fig. 1 LILBID spectrum of a 0.1 M solution of Na_3PO_4.

The $(NaPO_3)Na^+$ peak (125 u) cannot be reproduced with sufficient amplitude and the ratio of $(Na_3PO_4)Na^+$ (187 u) to $(Na_2HPO_4)Na^+$ (165 u) is too high compared with CDA spectra. Both these disagreements disappear if sufficient amounts of Na_2HPO_4 are added to the solution.

Extended Data Fig. 2 LILBID spectrum of a 0.1 M solution of Na_2HPO_4.

A mass line at 143 u ($(NaH_2PO_4)Na^+$) appears, which is not apparent in CDA spectra except tentatively in spectrum #3 of Extended Data Table 2. This signature at 143 u is not present in a mixture of Na_2HPO_4 and Na_3PO_4 (Fig. 2). This motivates our conclusion on a mixture of the two orthophosphates in the ice grains from Enceladus analysed by CDA.

Extended Data Fig. 3 LILBID spectrum with a 0.5 M phosphate concentration (similar to Fig. 2) and 0.01 M phosphite (Na_2HPO_3).

The phosphite peak $Na_2HPO_3Na^+$ (149 u) is not in agreement with CDA spectra and only becomes undetectable in phosphate-bearing CDA spectra at phosphite concentrations below 0.005 M. Note, the CDA detection limit of phosphite in pure water (without phosphates) is approx. 1 nM, more than 3 orders of magnitude lower.

Extended Data Fig. 4 Jaguar quantum chemistry program package at the PBE0-D3/6-311+G(d,p)/PBE level of theory of the elimination reaction from dihydrogen phosphate leading to the observed meta phosphate species.

The presence of water significantly reduces the meta phosphate formation barrier, specifically in the case with Na^+.

Extended Data Fig. 5 Typical Backscattered Electrons (BSE) SEM images and elemental mapping for solid residues collected after experiment Run #1.

The presence of large Ca carbonate grains (>100 µm) indicates the formation of calcite during the experiment. Such large Ca carbonate grains were absent from the sample before our experiments (Extended Data Fig. 7). XRD analysis (Extended Data Fig. 10) shows that the precipitated Ca carbonates are most likely calcite.

Extended Data Fig. 6 (a) Typical Backscattered Electrons (BSE) SEM images and elemental mapping results for Ca phosphate minerals in the solid residues collected after experiment Run #1.

Phosphate-rich grains with a size of several micrometres and veins with a width of several micrometres and length of ~100 µm are seen near a large serpentine grain. In addition to Ca and P, Na is enriched in Ca phosphate minerals. Close-up images of phosphate minerals (the bottom two images) indicate that Ca phosphate coexists with other minerals (e.g., serpentine and saponite). **(b)** Same as Extended Data Fig. 6a, but for experiment Run #2. Phosphate-rich grains with a size of several micrometres are observed (see

the bottom two SEM images for close-up views of phosphate minerals). Close-up images of Grain D indicate that Ca phosphate minerals coexist with other minerals, whereas the surface of Grain C exhibits relatively uniform composition. The chemical compositions of Ca phosphate (Grain A and Vein B) revealed by microprobe spot analysis are summarized in Extended Data Table 4. XRD analysis (Fig. 10) shows that the precipitated Ca carbonates are most likely calcite.

Extended Data Fig. 7 Typical SEM images and elemental mapping results of carbonaceous chondrite samples before the experiments.

The size of Ca carbonate grains is typically ~10 μm or less.

Extended Data Fig. 8 Temporal variations in the concentrations of selected elements (Na, K, Si, Mg, Ca, and P) and dissolved total inorganic carbon (ΣCO_2).

Results for (a) Run #1 and (b) Run #2 are shown. The arrows for Mg concentrations at 650 and 1000 h of reaction time for Run #2 indicate that the concentrations were less than the detection limit (0.036 mM). The measured data and calculated in situ pH of fluids are summarized in Extended Data Table 3. The decreases in ΣCO_2 in both experiments over time were caused by the formation of carbonate minerals (i.e., calcite), as confirmed by microprobe and the results of XRD analysis shown in Extended Data Figs. 5 and 10, respectively. Where no error bars are shown, the analytical errors are smaller than the data points. All measured data, analytical errors, and calculated in situ pH of fluids are summarized in Extended Data Table 3.

Extended Data Fig. 9 Ternary diagram of the results of EPMA analysis of the chemical compositions of phyllosilicate and Ca phosphate-rich grains in the samples before and after the experiments, plotted as molar fractions.

Green stars indicate the compositions of pure minerals found in the samples.

Extended Data Fig. 10 Comparison of X-ray diffraction (XRD) spectra of the solid samples before the experiments (red), after experiment Run #1 (purple), and after experiment Run #2 (blue). a.u., arbitrary units.

The XRD intensities are offset for clarity. The primary XRD peaks of minerals that correspond to the XRD peaks are also indicated.

Extended Data Fig. 11 Results of calculated ΣPO_4^{3-} in fluids as a function of temperature for different pH and ΣCO_2 assuming the whitlockite–calcite buffer system.

Pink shading represents the observationally-derived range of ΣPO_4^{3-} from the present study.

Extended Data Table 1 CDA set of Type 3 spectra used for this work
Extended Data Table 2 Events of phosphate-rich ice grain recorded by CDA. Saturn distance is given in Saturn radii (equatorial radius R_S = 60268 km)
Extended Data Table 3 Temporal variations in dissolved species in fluids (in M), with pH at the calculated in-situ pH at 150 °C and 30 MPa
Extended Data Table 4 Results (mole abundance in mol/kg) of electron probe microanalyzer (EPMA) analyses for Ca phosphates [Grain A and Vein B for Run #1, and Grains C and D for Run #2 (see Extended Data Fig. 6)], and for Ca carbonate minerals for Runs #1 and #2

Rights and permissions

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The

images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by/4.0/.

Reprints and Permissions

About this article

Cite this article

Postberg, F., Sekine, Y., Klenner, F. *et al.* Detection of phosphates originating from Enceladus's ocean. *Nature* **618**, 489–493 (2023). https://doi.org/10.1038/s41586-023-05987-9

- Received: 26 July 2022
- Accepted: 21 March 2023
- Published: 14 June 2023
- Issue Date: 15 June 2023
- DOI: https://doi.org/10.1038/s41586-023-05987-9

Share this article

Anyone you share the following link with will be able to read this content:

Get shareable link

Sorry, a shareable link is not currently available for this article.

Copy to clipboard

Provided by the Springer Nature SharedIt content-sharing initiative

Phosphate discovery hints at geochemistry and origin of Enceladus

- Mikhail Yu. Zolotov

Nature News & Views 14 Jun 2023

This article was downloaded by **calibre** from https://www.nature.com/articles/s41586-023-05987-9

| Section menu | Main menu |

- Article
- Open Access
- Published: 14 June 2023

Evidence for the utility of quantum computing before fault tolerance

- Youngseok Kim ,
- Andrew Eddins ,
- Sajant Anand,
- Ken Xuan Wei,
- Ewout van den Berg,
- Sami Rosenblatt,
- Hasan Nayfeh,
- Yantao Wu,
- Michael Zaletel,
- Kristan Temme &
- …
- Abhinav Kandala

Show authors

Nature **volume 618**, pages 500–505 (2023)

- 33k Accesses

- 385 Altmetric

- Metrics details

Abstract

Quantum computing promises to offer substantial speed-ups over its classical counterpart for certain problems. However, the greatest impediment to realizing its full potential is noise that is inherent to these systems. The widely accepted solution to this challenge is the implementation of fault-tolerant quantum circuits, which is out of reach for current processors. Here we report experiments on a noisy 127-qubit processor and demonstrate the measurement of accurate expectation values for circuit volumes at a scale beyond brute-force classical computation. We argue that this represents evidence for the utility of quantum computing in a pre-fault-tolerant era. These experimental results are enabled by advances in the coherence and calibration of a superconducting processor at this scale and the ability to characterize[1] and controllably manipulate noise across such a large device. We establish the accuracy of the measured expectation values by comparing them with the output of exactly verifiable circuits. In the regime of strong entanglement, the quantum computer provides correct results for which leading classical approximations such as pure-state-based 1D (matrix product states, MPS) and 2D (isometric tensor network states, isoTNS) tensor network methods[2,3] break down. These experiments demonstrate a foundational tool for the realization of near-term quantum applications[4,5].

Main

It is almost universally accepted that advanced quantum algorithms such as factoring[6] or phase estimation[7] will require quantum error correction. However, it is acutely debated whether processors available at present can be made sufficiently reliable to run other, shorter-depth quantum circuits at a scale that could provide an advantage for practical problems. At this point, the conventional expectation is that the implementation of even simple quantum circuits with the potential to exceed classical capabilities will have to wait until more advanced, fault-tolerant processors arrive. Despite the tremendous progress of quantum hardware in recent years, simple fidelity bounds[8] support this bleak forecast; one estimates that a quantum circuit 100 qubits wide by 100 gate-layers deep executed with 0.1% gate error yields a state fidelity less than 5×10^{-4}. Nonetheless, the question remains whether properties of the ideal state can be accessed even with such low fidelities. The error-mitigation[9,10] approach to near-term quantum advantage on noisy

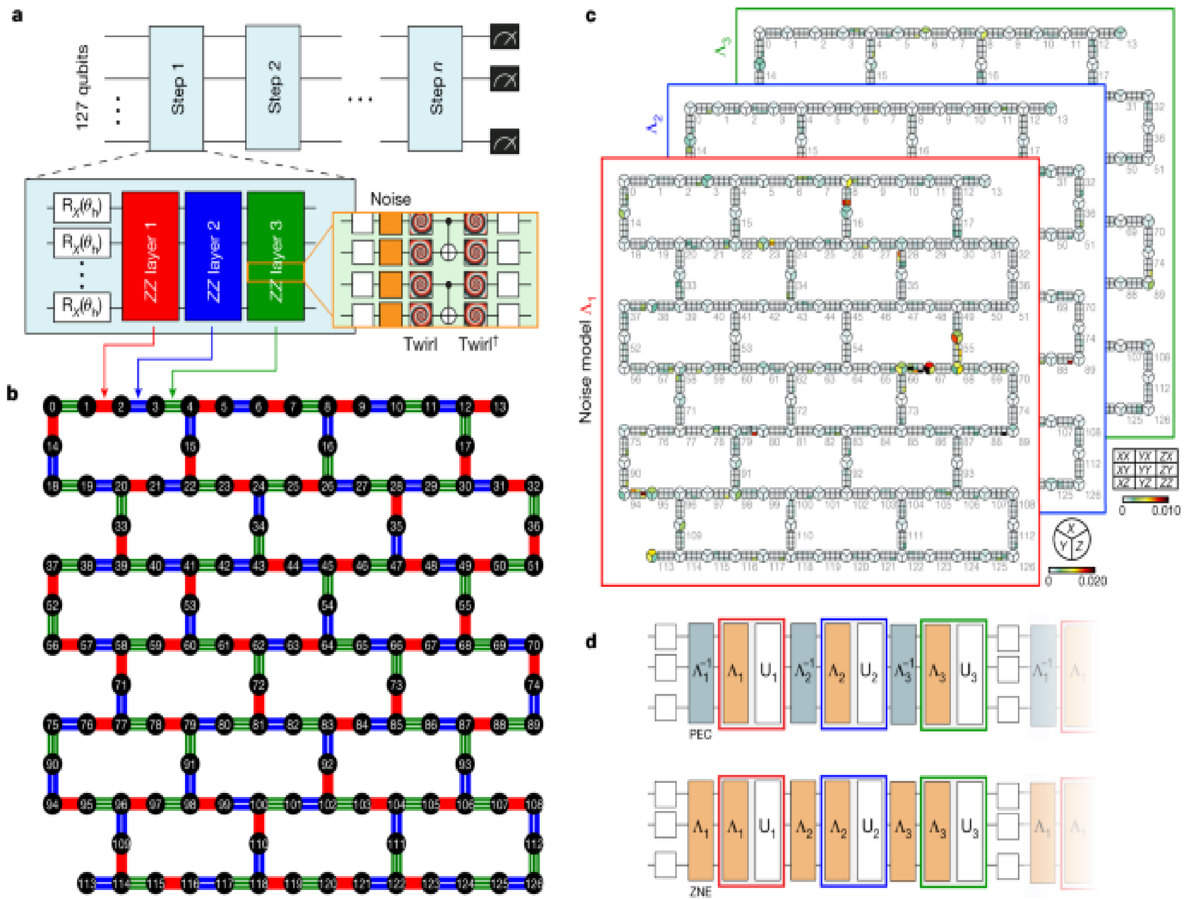
devices exactly addresses this question, that is, that one can produce accurate expectation values from several different runs of the noisy quantum circuit using classical post-processing.

Quantum advantage can be approached in two steps: first, by demonstrating the ability of existing devices to perform accurate computations at a scale that lies beyond brute-force classical simulation, and second by finding problems with associated quantum circuits that derive an advantage from these devices. Here we focus on taking the first step and do not aim to implement quantum circuits for problems with proven speed-ups.

We use a superconducting quantum processor with 127 qubits to run quantum circuits with up to 60 layers of two-qubit gates, a total of 2,880 CNOT gates. General quantum circuits of this size lie beyond what is feasible with brute-force classical methods. We thus first focus on specific test cases of the circuits permitting exact classical verification of the measured expectation values. We then turn to circuit regimes and observables in which classical simulation becomes challenging and compare with results from state-of-the-art approximate classical methods.

Our benchmark circuit is the Trotterized time evolution of a 2D transverse-field Ising model, sharing the topology of the qubit processor (Fig. 1a). The Ising model appears extensively across several areas in physics and has found creative extensions in recent simulations exploring quantum many-body phenomena, such as time crystals[11,12], quantum scars[13] and Majorana edge modes[14]. As a test of utility of quantum computation, however, the time evolution of the 2D transverse-field Ising model is most relevant in the limit of large entanglement growth in which scalable classical approximations struggle.

Fig. 1: Noise characterization and scaling for 127-qubit Trotterized time-evolution circuits.



a, Each Trotter step of the Ising simulation includes single-qubit X and two-qubit ZZ rotations. Random Pauli gates are inserted to twirl (spirals) and controllably scale the noise of each CNOT layer. The dagger indicates conjugation by the ideal layer. **b**, Three depth-1 layers of CNOT gates suffice to realize interactions between all neighbour pairs on ibm_kyiv. **c**, Characterization experiments efficiently learn the local Pauli error rates $\lambda_{l,i}$ (colour scales) comprising the overall Pauli channel Λ_l associated with the lth twirled CNOT layer. (Figure expanded in Supplementary Information IV.A). **d**, Pauli errors inserted at proportional rates can be used to either cancel (PEC) or amplify (ZNE) the intrinsic noise.

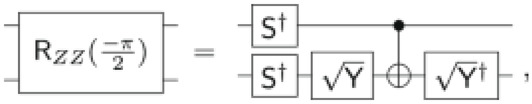
In particular, we consider time dynamics of the Hamiltonian,

$$H=-\,J\sum _{\langle i,j\rangle }{Z}_{i}{Z}_{j}+h\sum _{i}{X}_{i},$$
(1)

in which $J > 0$ is the coupling of nearest-neighbour spins with $i < j$ and h is the global transverse field. Spin dynamics from an initial state can be simulated by means of first-order Trotter decomposition of the time-evolution operator,

$$\begin{array}{ll}{\rm e}^{-{\rm i}H_{ZZ}\delta t}\,= & \prod_{\langle i,j\rangle}{\rm e}^{{\rm i}J\delta t Z_{i}Z_{j}}=\prod_{\langle i,j\rangle}{\rm R}_{Z_{i}Z_{j}}(-2J\delta t)\\ {\rm e}^{-{\rm i}H_{X}\delta t}\,= & \prod_{i}{\rm e}^{-{\rm i}h\delta t X_{i}}=\prod_{i}{\rm R}_{X_{i}}(2h\delta t),\\ \end{array}$$
(2)

in which the evolution time T is discretized into $T/\delta t$ Trotter steps and ${\rm R}_{Z_i Z_j}(\theta_J)$ and ${\rm R}_{X_i}(\theta_{\rm h})$ are ZZ and X rotation gates, respectively. We are not concerned with the model error owing to Trotterization and thus take the Trotterized circuit as ideal for any classical comparison. For experimental simplicity, we focus on the case $\theta_J = -2J\delta t = -\pi/2$ such that the ZZ rotation requires only one CNOT,



where the equality holds up to a global phase. In the resulting circuit (Fig. 1a), each Trotter step amounts to a layer of single-qubit rotations, ${\rm R}_X(\theta_{\rm h})$, followed by commuting layers of parallelized two-qubit rotations, ${\rm R}_{ZZ}(\theta_J)$.

For the experimental implementation, we primarily used the IBM Eagle processor ibm_kyiv, composed of 127 fixed-frequency transmon qubits[15] with heavy-hex connectivity and median T_1 and T_2 times of 288 μs and 127 μs, respectively. These coherence times are unprecedented for superconducting processors of this scale and allow the circuit depths accessed in this work. The two-qubit CNOT gates between neighbours are realized by calibrating the cross-resonance interaction[16]. As each qubit has at most three neighbours, all ZZ interactions can be performed in three layers of parallelized CNOT gates (Fig. 1b). The CNOT gates within each layer are

calibrated for optimal simultaneous operation (see Methods for more details).

We now see that these hardware performance improvements enable even larger problems to be successfully executed with error mitigation, in comparison with recent work[1,17] on this platform. Probabilistic error cancellation (PEC)[9] has been shown[1] to be very effective at providing unbiased estimates of observables. In PEC, a representative noise model is learned and effectively inverted by sampling from a distribution of noisy circuits related to the learned model. Yet, for the current error rates on our device, the sampling overhead for the circuit volumes considered in this work remains restrictive, as discussed further below.

We therefore turn to zero-noise extrapolation (ZNE)[9,10,17,18], which provides a biased estimator at a potentially much lower sampling cost. ZNE is either a polynomial[9,10] or exponential[19] extrapolation method for noisy expectation values as a function of a noise parameter. This requires the controlled amplification of the intrinsic hardware noise by a known gain factor G to extrapolate to the ideal $G = 0$ result. ZNE has been widely adopted in part because noise-amplification schemes based on pulse stretching[9,17,18] or subcircuit repetition[20,21,22] have circumvented the need for precise noise learning, while relying on simplistic assumptions about the device noise. More precise noise amplification can, however, enable substantial reductions in the bias of the extrapolated estimator, as we demonstrate here.

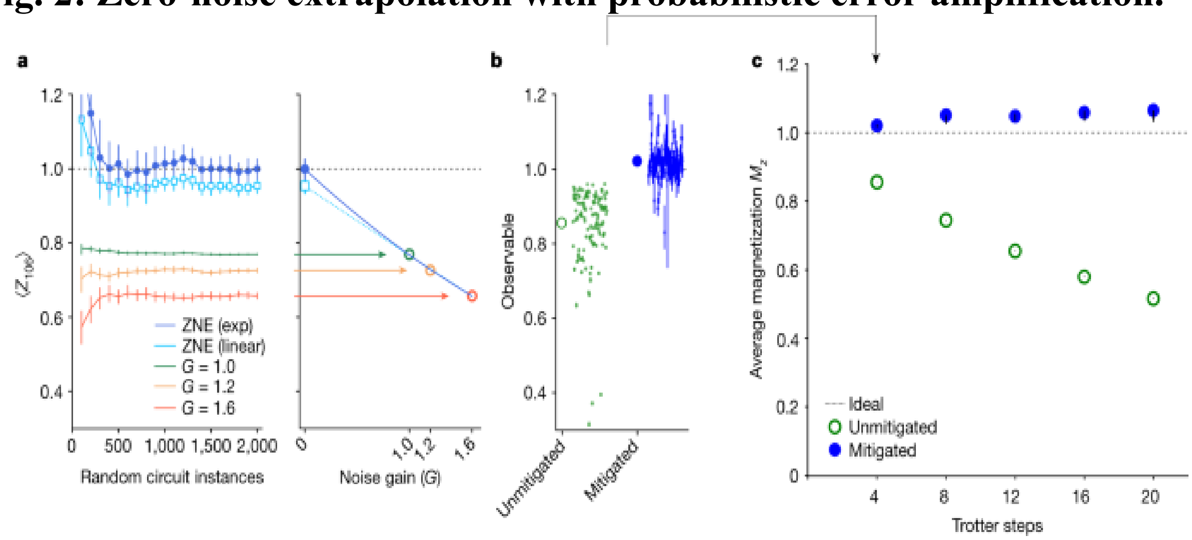
The sparse Pauli–Lindblad noise model proposed in ref. [1] turns out to be especially well suited for noise shaping in ZNE. The model takes the form ${{\rm{e}}}^{{\mathcal{L}}}$, in which ${\mathcal{L}}$ is a Lindbladian comprising Pauli jump operators P_i weighted by rates λ_i. It was shown in ref. [1] that restricting to jump operators acting on local pairs of qubits yields a sparse noise model that can be efficiently learned for many qubits and that accurately captures the noise associated with layers of two-qubit Clifford gates, including crosstalk, when combined with random Pauli twirls[23,24]. The noisy layer of gates is modelled as a set of ideal gates preceded by some noise channel Λ. Thus, applying Λ^{α} before the noisy layer produces an overall noise channel Λ^{G} with gain $G = \alpha + 1$. Given the exponential form of the Pauli–Lindblad noise model, the map \

({{\rm{e}}}^{\alpha {\mathcal{L}}}\) is obtained by simply multiplying the Pauli rates λ_i by α. The resulting Pauli map can be sampled to obtain appropriate circuit instances; for $\alpha \geq 0$, the map is a Pauli channel that can be sampled directly, whereas for $\alpha < 0$, quasi-probabilistic sampling is needed with sampling overhead $\gamma^{-2\alpha}$ for some model-specific γ. In PEC, we choose $\alpha = -1$ to obtain an overall zero-gain noise level. In ZNE, we instead amplify the noise[10,25,26,27] to different gain levels and estimate the zero-noise limit using extrapolation. For practical applications, we need to consider the stability of the learned noise model over time (Supplementary Information III.A), for instance, owing to qubit interactions with fluctuating microscopic defects known as two-level systems[28].

Clifford circuits serve as useful benchmarks of estimates produced by error mitigation, as they can be efficiently simulated classically[29]. Notably, the entire Ising Trotter circuit becomes Clifford when θ_{h} is chosen to be a multiple of π/2. As a first example, we therefore set the transverse field to zero ($\mathrm{R}_X(0) = I$) and evolve the initial state $|0\rangle^{\otimes 127}$ (Fig. 1a). The CNOT gates nominally leave this state unchanged, so the ideal weight-1 observables Z_q all have expectation value 1; owing to the Pauli twirling of each layer, the bare CNOTs do affect the state. For each Trotter experiment, we first characterized the noise models Λ_l for the three Pauli-twirled CNOT layers (Fig. 1c) and then used these models to implement Trotter circuits with noise gain levels $G \in \{1, 1.2, 1.6\}$. Figure 2a illustrates the estimation of $\langle Z_{106}\rangle$ after four Trotter steps (12 CNOT layers). For each G, we generated 2,000 circuit instances in which, before each layer l, we have inserted products of one-qubit and two-qubit Pauli errors i from \({\mathcal{L}}\) drawn with probabilities \({p}_{l,i}=(1-{{\rm{e}}}^{-2(G-1){\lambda }_{l,i}})/2\) and executed each instance 64 times, totalling 384,000 executions. As more circuit instances are accumulated, the estimates of $\langle Z_{106}\rangle_G$, corresponding to the different gains G, converge to distinct values. The different estimates are then fit by an extrapolating function in G to estimate the ideal value $\langle Z_{106}\rangle_0$. The results in Fig. 2a highlight the reduced bias from exponential extrapolation[19] in comparison with linear extrapolation. That said, exponential extrapolation can exhibit instabilities, for instance, when expectation values are

unresolvably close to zero, and—in such cases—we iteratively downgrade the extrapolation model complexity (see Supplementary Information II.B). The procedure outlined in Fig. 2a was applied to the measurement results from each qubit q to estimate all $N = 127$ Pauli expectations $\langle Z_q \rangle_0$. The variation in the unmitigated and mitigated observables in Fig. 2b is indicative of the non-uniformity in the error rates across the entire processor. We report the global magnetization along $\hat{z}$, ${M}_{z}={\sum }_{q}\,\langle {Z}_{q}\rangle /N$, for increasing depth in Fig. 2c. Although the unmitigated result shows a gradual decay from 1 with an increasing deviation for deeper circuits, ZNE greatly improves agreement, albeit with a small bias, with the ideal value even out to 20 Trotter steps, or 60 CNOT depth. Notably, the number of samples used here is much smaller than an estimate of the sampling overhead that would be needed in a naive PEC implementation (see Supplementary Information IV.B). In principle, this disparity may be greatly reduced by more advanced PEC implementations using light-cone tracing[30] or by improvements in hardware error rates. As future hardware and software developments bring down sampling costs, PEC may be preferred when affordable to avoid the potentially biased nature of ZNE.

Fig. 2: Zero-noise extrapolation with probabilistic error amplification.

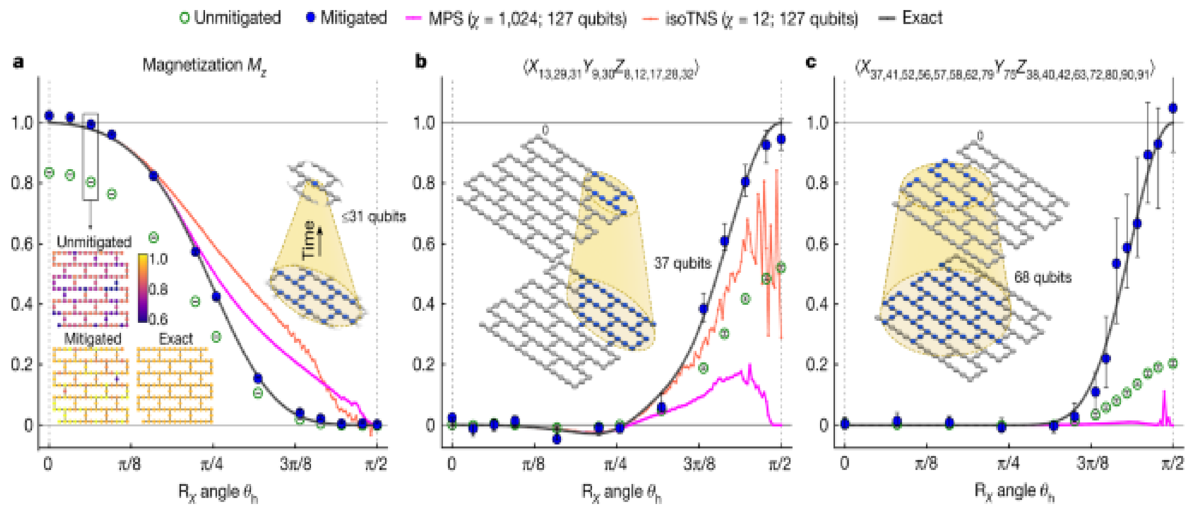


Mitigated expectation values from Trotter circuits at the Clifford condition $\theta_h = 0$. **a**, Convergence of unmitigated ($G = 1$), noise-amplified ($G > 1$) and noise-mitigated (ZNE) estimates of $\langle Z_{106} \rangle$ after four Trotter steps. In all panels, error bars indicate 68% confidence intervals obtained by means of

percentile bootstrap. Exponential extrapolation (exp, dark blue) tends to outperform linear extrapolation (linear, light blue) when differences between the converged estimates of $\langle Z_{106}\rangle_{G\neq 0}$ are well resolved. **b**, Magnetization (large markers) is computed as the mean of the individual estimates of $\langle Z_q\rangle$ for all qubits (small markers). **c**, As circuit depth is increased, unmitigated estimates of M_z decay monotonically from the ideal value of 1. ZNE greatly improves the estimates even after 20 Trotter steps (see Supplementary Information II for ZNE details).

Next, we test the efficacy of our methods for non-Clifford circuits and the Clifford $\theta_\mathrm{h} = \pi/2$ point, with non-trivial entangling dynamics compared with the identity-equivalent circuits discussed in Fig. 2. The non-Clifford circuits are of particular importance to test, as the validity of exponential extrapolation is no longer guaranteed (see Supplementary Information V and ref. [31]). We restrict the circuit depth to five Trotter steps (15 CNOT layers) and judiciously choose observables that are exactly verifiable. Figure 3 shows the results as θ_h is swept between 0 and π/2 for three such observables of increasing weight. Figure 3a shows M_z as before, an average of weight-1 $\langle Z\rangle$ observables, whereas Fig. 3b,c show weight-10 and weight-17 observables. The latter operators are stabilizers of the Clifford circuit at $\theta_\mathrm{h} = \pi/2$, obtained by evolution of the initial stabilizers Z_{13} and Z_{58}, respectively, of $|0\rangle^{\otimes 127}$ for five Trotter steps, ensuring non-vanishing expectation values in the strongly entangling regime of particular interest. Although the entire 127-qubit circuit is executed experimentally, light-cone and depth-reduced (LCDR) circuits enable brute-force classical simulation of the magnetization and weight-10 operator at this depth (see Supplementary Information VII). Over the full extent of the θ_h sweep, the error-mitigated observables show good agreement with the exact evolution (see Fig. 3a,b). However, for the weight-17 operator, the light cone expands to 68 qubits, a scale beyond brute-force classical simulation, so we turn to tensor network methods.

Fig. 3: Classically verifiable expectation values from 127-qubit, depth-15 Clifford and non-Clifford circuits.



Expectation value estimates for θ_h sweeps at a fixed depth of five Trotter steps for the circuit in Fig. 1a. The considered circuits are non-Clifford except at $\theta_h = 0, \pi/2$. Light-cone and depth reductions of respective circuits enable exact classical simulation of the observables for all θ_h. For all three plotted quantities (panel titles), mitigated experimental results (blue) closely track the exact behaviour (grey). In all panels, error bars indicate 68% confidence intervals obtained by means of percentile bootstrap. The weight-10 and weight-17 observables in **b** and **c** are stabilizers of the circuit at $\theta_h = \pi/2$ with respective eigenvalues +1 and −1; all values in **c** have been negated for visual simplicity. The lower inset in **a** depicts variation of $\langle Z_q \rangle$ at $\theta_h = 0.2$ across the device before and after mitigation and compares with exact results. Upper insets in all panels illustrate causal light cones, indicating in blue the final qubits measured (top) and the nominal set of initial qubits that can influence the state of the final qubits (bottom). M_z also depends on 126 other cones besides the example shown. Although in all panels exact results are obtained from simulations of only causal qubits, we include tensor network simulations of all 127 qubits (MPS, isoTNS) to help gauge the domain of validity for those techniques, as discussed in the main text. isoTNS results for the weight-17 operator in **c** are not accessible with current methods (see Supplementary Information VI). All experiments were carried out for $G = 1, 1.2, 1.6$ and extrapolated as in Supplementary Information II.B. For each G, we generated 1,800–2,000 random circuit instances for **a** and **b** and 2,500–3,000 instances for **c**.

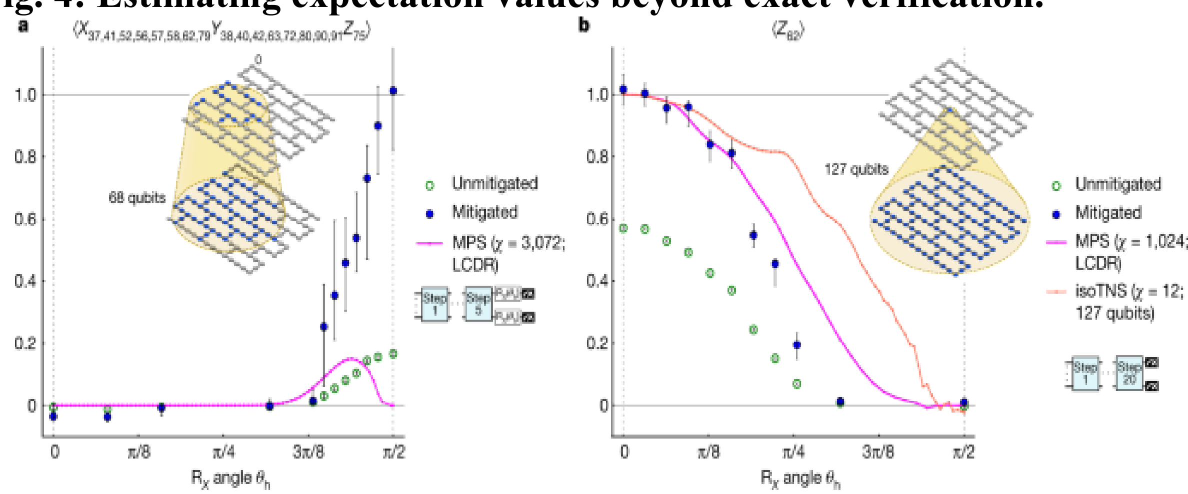
Tensor networks have been widely used to approximate and compress quantum state vectors that arise in the study of the low-energy eigenstates of and time evolution by local Hamiltonians[2,32,33] and, more recently, have been successfully used to simulate low-depth noisy quantum circuits[34,35,36]. Simulation accuracy can be improved by increasing the bond dimension χ, which constrains the amount of entanglement of the represented quantum state, at a computational cost scaling polynomially with χ. As entanglement (bond dimension) of a generic state grows linearly (exponentially) with time evolution until it saturates the volume law, deep quantum circuits are inherently difficult for tensor networks[37]. We consider both quasi-1D matrix product states (MPS) and 2D isometric tensor network states (isoTNS)[3] that have ${\mathcal{O}}({\chi }^{3})$ and ${\mathcal{O}}({\chi }^{7})$ scaling of time-evolution complexity, respectively. Details of both methods and their strengths are provided in Methods and Supplementary Information VI. Specifically for the case of the weight-17 operator shown in Fig. 3c, we find that an MPS simulation of the LCDR circuit at $\chi = 2{,}048$ is sufficient to obtain the exact evolution (see Supplementary Information VIII). The larger causal cone of the weight-17 observable results in an experimental signal that is weaker compared with that of the weight-10 observable; nevertheless, mitigation still yields good agreement with the exact trace. This comparison suggests that the domain of experimental accuracy could extend beyond the scale of exact classical simulation.

We expect that these experiments will eventually extend to circuit volumes and observables in which such light-cone and depth reductions are no longer important. Therefore, we also study the performance of MPS and isoTNS for the full 127-qubit circuit executed in Fig. 3, at respective bond dimensions of $\chi = 1{,}024$ and $\chi = 12$, which are primarily limited by memory requirements. Figure 3 shows that the tensor network methods struggle with increasing θ_{h}, losing both accuracy and continuity near the verifiable Clifford point $\theta_{\mathrm{h}} = \pi/2$. This breakdown can be understood in terms of entanglement properties of the state. The stabilizer state produced by the circuit at $\theta_{\mathrm{h}} = \pi/2$ has an exactly flat bipartite entanglement spectrum, found from a Schmidt decomposition of a 1D ordering of the qubits. Thus, truncating states with small Schmidt weight—the basis of all tensor network algorithms—is not justified. However, as exact tensor network

representations generically require bond dimension exponential in circuit depth, truncation is necessary for tractable numerical simulations.

Finally, in Fig. 4, we stretch our experiments to regimes in which the exact solution is not available with the classical methods considered here. The first example (Fig. 4a) is similar to Fig. 3c but with a further final layer of single-qubit Pauli rotations that interrupt the circuit-depth reduction that previously enabled exact verification for any θ_h (see Supplementary Information VII). At the verifiable Clifford point $\theta_h = \pi/2$, the mitigated results agree again with the ideal value, whereas the $\chi = 3{,}072$ MPS simulation of the 68-qubit LCDR circuit markedly fails in the strongly entangling regime of interest. Although $\chi = 2{,}048$ was sufficient for exact simulation of the weight-17 operator in Fig. 3c, an MPS bond dimension of 32,768 would be needed for exact simulation of this modified circuit and operator with $\theta_h = \pi/2$.

Fig. 4: Estimating expectation values beyond exact verification.



Plot markers, confidence intervals and causal light cones appear as defined in Fig. 3. **a**, Estimates of a weight-17 observable (panel title) after five Trotter steps for several values of θ_h. The circuit is similar to that in Fig. 3c but with further single-qubit rotations at the end. This effectively simulates the time evolution of the spins after Trotter step six by using the same number of two-qubit gates used for Trotter step five. As in Fig. 3c, the observable is a stabilizer at $\theta_h = \pi/2$ with eigenvalue −1, so we negate the *y* axis for visual simplicity. Optimization of the MPS simulation by including only qubits and gates in the causal light cone enables a higher bond dimension ($\chi = 3{,}072$), but the simulation still fails to approach −1 (+1 in

negated y axis) at $\theta_h = \pi/2$. **b**, Estimates of the single-site magnetization $\langle Z_{62}\rangle$ after 20 Trotter steps for several values of θ_h. The MPS simulation is light-cone-optimized and performed with bond dimension $\chi = 1{,}024$, whereas the isoTNS simulation ($\chi = 12$) includes the gates outside the light cone. The experiments were carried out with $G = 1, 1.3, 1.6$ for **a** and $G = 1, 1.2, 1.6$ for **b**, and extrapolated as in Supplementary Information II.B. For each G, we generated 2,000–3,200 random circuit instances for **a** and 1,700–2,400 instances for **b**.

As a final example, we extend the circuit depth to 20 Trotter steps (60 CNOT layers) and estimate the θ_h dependence of a weight-1 observable, $\langle Z_{62}\rangle$, in Fig. 4b, in which the causal cone extends over the entire device. Given the non-uniformity of device performance, also seen in the spread of single-site observables in Fig. 2b, we choose an observable that obtains the expected result $\langle Z_{62}\rangle \approx 1$ at the verifiable $\theta_h = 0$ point. Despite the greater depth, the MPS simulations of the LCDR circuit agree well with the experiment in the weakly entangling regime of small θ_h. Although deviations from the experimental trace emerge with increasing θ_h, we note that the MPS simulations slowly move in the direction of the experimental data with increasing χ (see Supplementary Information X) and that the bond dimension needed to exactly represent the stabilizer state and its evolution to depth 20 at $\theta_h = \pi/2$ is 7.2×10^{16}, 13 orders of magnitude larger than what we considered (see Supplementary Information VIII). For reference, as the memory required to store an MPS scales as \({\mathcal{O}}({\chi }^{2})\), already a bond dimension of $\chi = 1 \times 10^8$ would require 400 PB, independent of any runtime considerations. Furthermore, full-state tensor network simulations are already unable to capture the dynamics at the exactly verifiable five-step circuit in Fig. 3a. We also note that, given the large unmitigated signal, there may be opportunity to study time evolution at even larger depths on the current device.

For execution times, the tensor network simulations in Fig. 4 were run on a 64-core, 2.45-GHz processor with 128 GB of memory, in which the run time to access an individual data point at fixed θ_h was 8 h for Fig. 4a and 30 h for Fig. 4b. The corresponding quantum wall-clock run time was approximately 4 h for Fig. 4a and 9.5 h for Fig. 4b, but this is also far from a fundamental

limit, being at present dominated by classical processing delays that stand to be largely eliminated through conceptually straightforward optimizations. Indeed, the estimated device run time for the error-mitigated expectation values using 614,400 samples (2,400 circuit instances for each gain factor and readout error mitigation, with 64 shots per instance) at a conservative sampling rate of 2 kHz is only 5 min 7 s, which can be even further reduced by optimization of qubit reset speeds. On the other hand, the classical simulations may also be improved by methods besides the pure-state tensor networks considered here, such as Heisenberg operator evolution methods, which have recently been applied to non-Clifford simulations[38]. Another approach is to numerically emulate the ZNE used experimentally. For example, it was recently argued that the finite-χ truncation error introduced by tensor-product compression mimics experimental gate errors[34]. It would thus be natural to develop a theory for extrapolating tensor network state expectation values in the bond dimension χ for time evolution, as has been done in the case of ground-state search[39]. Alternatively, one can more directly emulate ZNE by introducing artificial dissipation into the dynamics engineered so that the resulting mixed-state evolution has reduced tensor-product bond dimension, as—for example—in dissipation-assisted operator evolution[40], and extrapolate results with respect to the strength of the dissipation. Although such methods[40,41] can successfully capture the long-time dynamics of the low-weight observables of a 1D spin chain, their applicability to high-weight observables in 2D at intermediate times is not clear—particularly as these methods are explicitly constructed to truncate complex operators.

The observation that a noisy quantum processor, even before the advent of fault-tolerant quantum computing, produces reliable expectation values at a scale beyond 100 qubits and non-trivial circuit depth leads to the conclusion that there is indeed merit to pursuing research towards deriving a practical computational advantage from noise-limited quantum circuits. Over recent years, substantial research effort has been directed to develop and demonstrate candidate heuristic quantum algorithms[5] that use noise-limited quantum circuits to estimate expectation values. We have now reached reliability at a scale for which one will be able to verify proposals and explore new approaches to determine which can provide utility beyond classical approximation methods. At the same time, these results will

motivate and help advance classical approximation methods as both approaches serve as valuable benchmarks of one another. However, even with improved classical methods, impending order-of-magnitude improvements in gate fidelities[42] and speed of superconducting quantum systems will drive substantial enhancements in accessible circuit volumes and paint an increasingly bright picture of the utility of noisy quantum computers.

Methods

Device calibration

The speed of cross-resonance-based CNOT gates is dependent on the qubit–qubit detuning and, typically, gate speeds across the device are chosen independently to minimize individual gate errors[43]. This leads to a large spread in CNOT times across the device. Noting that the speed of each parallelized CNOT layer is limited by the slowest gate in the layer, we develop a new tune-up scheme for large-scale processor calibration that optimizes the layer rather than the individual gates. First, the control and target qubits are assigned to each gate layer to reduce crosstalk and leakage from transmon-frequency collisions. The slowest gate in each layer then has its duration carefully optimized. Finally, all gates in the layer are fixed to this duration and calibrated simultaneously with error-amplification sequences[44]. Compared with independently calibrated gates, the layer duration is unchanged, but gates are slower with lower drive amplitudes, reducing any leakage arising from multi-photon transitions. The simultaneous calibration also ensures that the gates are calibrated as they are implemented in the circuit.

Noise model

Throughout this work, we amplify gate noise by means of a learned noise model. For this model, following ref. [1], a general Pauli channel is approximated by $\Lambda (\rho)=\exp [{\mathcal{L}}](\rho)$ with a sparse Pauli–Lindblad generator

$${\mathcal{L}}(\rho)=\sum _{i}{\lambda }_{i}({P}_{i}\,\rho {P}_{i}^{\dagger }-\rho).$$

Here the jump operators are chosen to be Pauli operators P_i with $({P}_{i}^{\dagger }{P}_{i}=I)$ and the model is parameterized by the non-negative coefficients λ_i. This model can be rewritten as

$$\Lambda (\rho)=\mathop{\bigcirc}\limits_{i}\left({w}_{i}\cdot +(1-{w}_{i}){P}_{i}\cdot {P}_{i}^{\dagger }\right)(\rho),$$

in which ${w}_{i}=(1+{{\rm{e}}}^{-2{\lambda }_{i}})/2$ and ${\bigcirc}_{i=1}^{n}{O}_{i}(\,\cdot \,)=({O}_{n}\,\circ \,{O}_{n-1}\,\circ \cdots \circ \,{O}_{1})$ represents the composition of operators and $O(\cdot)(\rho) = O(\rho)$. In other words, we can express $\Lambda(\rho)$ as a composition of simple Pauli maps. For physical noise channels, in which all $\lambda_i \geq 0$, the composition consists of simply Pauli channels. By allowing non-zero coefficients λ_i only for Pauli terms P_i whose support corresponds to a single qubit or a pair of connected qubits, we obtain a sparse noise model that can be efficiently learned and that, despite its simplicity, is able to capture crosstalk errors[1]. It is readily seen that $\exp [\alpha {\mathcal{L}}]$ is obtained by scaling all λ_i by α. For $\alpha \geq 0$, the resulting noise model is a composition of Pauli channels. Samples from this channel can be obtained by independently sampling P_i with probability $1 - w_i$ for each of the simple channels and multiplying the results. For $\alpha < 0$, the resulting coefficients $1 - w_i$ are generally negative, leading to a non-physical noise map. Sampling in that case can still be done, albeit in a quasi-probabilistic manner. Doing so results in a sampling overhead of γ^2, in which $\gamma =\exp \left({\sum }_{i}2{\lambda }_{i}\right)$.

Brute-force simulations

The simplest, most accurate and most limited method is simulation of a collection of the state of M qubits as a dense vector of 2^M complex coefficients. All unitary gates, irrespective of locality, can be applied directly to the vector. Expectation values are found by vector–matrix–vector product

of the conjugated state, operator and state. We use this approach for simulations up to 30 qubits.

Tensor network methods

For circuits of more than 30 qubits, we used 1D and 2D tensor network state methods[45]. For a quantum state on M qubits, tensor network methods approximate the 2^M complex coefficients for the wavefunction amplitude as a network of contracted tensors containing ${\mathcal{O}}\left(M{\chi }^{p}\right)$ coefficients, in which p is an integer depending on the method. Here we consider MPS[2,32,33] with $p = 2$ and isoTNS[3] with $p = 4$. MPS represent a quantum state as a network of rank-3 tensors that, when contracted or multiplied together, give an approximation to the wavefunction amplitude for each basis state. isoTNS are a restriction of projected entangled pair states, a 2D generalization of MPS to square lattices in which the network consists of rank-5 tensors. The accuracy and computational cost of both MPS and isoTNS depend on the bond dimension χ. MPS methods have the advantage of well-developed algorithms, yet suffer from fundamental limitations of using a 1D method to simulate a 2D system. isoTNS methods, on the other hand, are inherently 2D methods but suffer from unavoidable sources of error not present for MPS, though these can be systematically reduced with increasing bond dimension.

Data availability

The datasets generated and analysed during this study are available at https://doi.org/10.6084/m9.figshare.22500355.

References

1. van den Berg, E., Minev, Z.K., Kandala, A. et al. Probabilistic error cancellation with sparse Pauli–Lindblad models on noisy quantum processors. *Nat. Phys.* https://doi.org/10.1038/s41567-023-02042-2 (2023).

2. Paeckel, S. et al. Time-evolution methods for matrix-product states. *Ann. Phys.* **411**, 167998 (2019).

3. Zaletel, M. P. & Pollmann, F. Isometric tensor network states in two dimensions. *Phys. Rev. Lett.* **124**, 037201 (2020).

4. Preskill, j Quantum computing in the NISQ era and beyond. *Quantum* **2**, 79 (2018).

5. Bharti, K. et al. Noisy intermediate-scale quantum algorithms. *Rev. Mod. Phys.* **94**, 015004 (2022).

6. Shor, P. W. in *Proc. 35th Annual Symposium on Foundations of Computer Science* 124–134 (IEEE, 1994).

7. Kitaev, A. Y. Quantum measurements and the Abelian Stabilizer Problem. Preprint at https://arxiv.org/abs/quant-ph/9511026 (1995).

8. Arute, F. et al. Quantum supremacy using a programmable superconducting processor. *Nature* **574**, 505–510 (2019).

9. Temme, K., Bravyi, S. & Gambetta, J. M. Error mitigation for short-depth quantum circuits. *Phys. Rev. Lett.* **119**, 180509 (2017).

10. Li, Y. & Benjamin, S. C. Efficient variational quantum simulator incorporating active error minimization. *Phys. Rev. X* **7**, 021050 (2017).

11. Mi, X. et al. Time-crystalline eigenstate order on a quantum processor. *Nature* **601**, 531–536 (2022).

12. Frey, P. & Rachel, S. Realization of a discrete time crystal on 57 qubits of a quantum computer. *Sci. Adv.* **8**, eabm7652 (2022).

13. Chen, I.-C., Burdick, B., Yao, Y., Orth, P. P. & Iadecola, T. Error-mitigated simulation of quantum many-body scars on quantum computers with pulse-level control. *Phys. Rev. Res.* **4**, 043027 (2022).

14. Mi, X. et al. Noise-resilient edge modes on a chain of superconducting qubits. *Science* **378**, 785–790 (2022).

15. Koch, J. et al. Charge-insensitive qubit design derived from the Cooper pair box. *Phys. Rev. A* **76**, 042319 (2007).

16. Chow, J. M. et al. Simple all-microwave entangling gate for fixed-frequency superconducting qubits. *Phys. Rev. Lett.* **107**, 080502 (2011).

17. Kim, Y. et al. Scalable error mitigation for noisy quantum circuits produces competitive expectation values. *Nat. Phys.* https://doi.org/10.1038/s41567-022-01914-3 (2023).

18. Kandala, A. et al. Error mitigation extends the computational reach of a noisy quantum processor. *Nature* **567**, 491–495 (2019).

19. Endo, S., Benjamin, S. C. & Li, Y. Practical quantum error mitigation for near-future applications. *Phys. Rev. X* **8**, 031027 (2018).

20. Dumitrescu, E. F. et al. Cloud quantum computing of an atomic nucleus. *Phys. Rev. Lett.* **120**, 210501 (2018).

21. He, A., Nachman, B., de Jong, W. A. & Bauer, C. W. Zero-noise extrapolation for quantum-gate error mitigation with identity insertions. *Phys. Rev. A* **102**, 012426 (2020).

22. Giurgica-Tiron, T., Hindy, Y., Larose, R., Mari, A. & Zeng, W. J. digital zero noise extrapolation for quantum error mitigation. in *2020 IEEE International Conference on Quantum Computing and Engineering (QCE)* 306–316 (IEEE, 2020).

23. Bennett, C. H. et al. Purification of noisy entanglement and faithful teleportation via noisy channels. *Phys. Rev. Lett.* **76**, 722–725 (1996).

24. Knill, E. Fault-tolerant postselected quantum computation: threshold analysis. Preprint at https://arxiv.org/abs/quant-ph/0404104 (2004).

25. Mari, A., Shammah, N. & Zeng, W. J. Extending quantum probabilistic error cancellation by noise scaling. *Phys. Rev. A* **104**, 052607 (2021).

26. Ferracin, S. et al. Efficiently improving the performance of noisy quantum computers. Preprint at https://arxiv.org/abs/2201.10672

(2022).

27. McDonough, B. et al. Automated quantum error mitigation based on probabilistic error reduction. In I*EEE/ACM Third International Workshop on Quantum Computing Software (QCS)* 83–93 (IEEE/ACM, 2022).

28. Carroll, M., Rosenblatt, S., Jurcevic, P., Iauer, I. & Kandala, A. Dynamics of superconducting qubit relaxation times. *npj Quantum Inf.* **9**, 132 (2022).

29. Aaronson, S. & Gottesman, D. Improved simulation of stabilizer circuits. *Phys. Rev. A* **70**, 052328 (2004).

30. Tran, M. C., Sharma, K. & Temme, K. Locality and error mitigation of quantum circuits. Preprint at https://arxiv.org/abs/2303.06496 (2023).

31. Cai, Z. Multi-exponential error extrapolation and combining error mitigation techniques for NISQ applications. *npj Quantum Inf.* **7**, 80 (2021).

32. Schöllwock, U. The density-matrix renormalization group in the age of matrix product states. *Ann. Phys.* **326**, 96–192 (2011).

33. Hauschild, J. & Pollmann, F. Efficient numerical simulations with tensor networks: Tensor Network Python (TeNPy). *SciPost Phys. Lect. Notes* https://doi.org/10.21468/SciPostPhysLectNotes.5 (2018).

34. Zhou, Y., Stoudenmire, E. M. & Waintal, X. What limits the simulation of quantum computers?. *Phys. Rev. X* **10**, 041038 (2020).

35. Guo, C. et al. General-purpose quantum circuit simulator with projected entangled-pair states and the quantum supremacy frontier. *Phys. Rev. Lett.* **123**, 190501 (2019).

36. Ayral, T. et al. Density-matrix renormalization group algorithm for simulating quantum circuits with a finite fidelity. *PRX Quantum* **4**, 020304 (2023).

37. Calabrese, P. & Cardy, J. Evolution of entanglement entropy in one-dimensional systems. *J. Stat. Mech. Theory Exp.* **2005**, P04010 (2005).

38. Bravyi, S. et al. Simulation of quantum circuits by low-rank stabilizer decompositions. *Quantum* **3**, 181 (2019).

39. Hubig, C., Haegeman, J. & Schollwöck, U. Error estimates for extrapolations with matrix-product states. *Phys. Rev. B* **97**, 045125 (2018).

40. Rakovszky, T., von Keyserlingk, C. W. & Pollmann, F. Dissipation-assisted operator evolution method for capturing hydrodynamic transport. *Phys. Rev. B* **105**, 075131 (2022).

41. White, C. D., Zaletel, M., Mong, R. S. K. & Refael, G. Quantum dynamics of thermalizing systems. *Phys. Rev. B* **97**, 035127 (2018).

42. Stehlik, J. et al. Tunable coupling architecture for fixed-frequency transmon superconducting qubits. *Phys. Rev. Lett.* **127**, 080505 (2021).

43. Zhang, E. J. et al. High-performance superconducting quantum processors via laser annealing of transmon qubits. *Sci. Adv.* **8**, eabi6690 (2022).

44. Wei, K. X. et al. Hamiltonian engineering with multicolor drives for fast entangling gates and quantum crosstalk cancellation. *Phys. Rev. Lett.* **129**, 060501 (2022).

45. Cirac, J. I., Pérez-García, D., Schuch, N. & Verstraete, F. Matrix product states and projected entangled pair states: concepts, symmetries, theorems. *Rev. Mod. Phys.* **93**, 045003 (2021).

Acknowledgements

This work would not have been possible without the contributions of the entire IBM Quantum team, including: K. Aarhus, B. Abdo, D. W. Abraham, E. Adams, G. Aleksandrowicz, D. Alevras, T. Alexander, A. Alexopoulos, H. Alghassi, L. Ament, M. Amico, A. Anderson, M. Aney, S. A. S.

Antezana, J. Apuzzo, E. Arbel, E. Arellano, V. Arena, T. Armon, M. Arthur, P. K. Austel, U. Bacher, S. Bangsaruntip, Z. Barabás, v. barbosa, S. V. Barron, G. Bauer, J. Bauer, M. Beck, M. Beckley, S. W. Bedell, L. Bello, Y. Ben-Haim, G. Bennett, L. A. Berge, M. Bernagozzi, J. Betke, L. S. Bishop, J. Blair, A. A. Blanco, S. R. Blanks, D. F. Bogorin, R. Bonam, M. Boraas, P. Bosavage, Y. Bosch, S. Bravyi, M. Brink, B. J. Brown, J. S. Broz, J. Bruley, D. Bryant, M. Buehler, M. Byers, M. T. Byrnes, E. Bäumer, C. Cabral Jr., S. Cairns, J. Calderón, L. Carata-Dejoianu, M. Carlucci, S. Carnevale, A. Carniol, S. Carri, M. S. Carroll, A. Carter, E. C. Castañeda, R. Chadwick, M. E. V. Chan, R. Cheek, V. Chiang, R. Chiarella, M. Chido, S. Choi, R. Cholarajan, A. Choquette, J. M. Chow, T. A. Christensen, B. Christenson, C. Christianson, W. Chung, C. Chung, M. Clausen, C. Codella, H. Colquhoun Jr., R. Conti, M. Copel, A. D. Corcoles, H. Cortes, J. M. Cotte, R. Cremins, A. Cross, A. Cross, G. Crossman, J. Cruz-Benito, S. Czaplewski, P. E. Dahlen, H. P. Dang, K. Dang, J. Dangler, M. P. David, A. R. Davila, G. Dawson, A. Dekusar, B. Delaney, P. Delosreyes, S. Derakhshandeh, A. Deshmukh, M. Desnoyers, O. E. Dial, M. DiDonato, S. Dikaleh, B. Dimock, T. Dixon, R. A. Donaton, B. Donovan, M. Doyle, Z. Dreiss, J. Drummond, E. A. Duch, N. Dunfee, L. Durst, R. A. Eades, N. Earnest-Noble, D. J. Egger, N. Einsidler, B. Ek, J. Ekman, S. Elrington, R. Elsasser, I. Elsayed, A. Emerich, S. Engelmann, K. R. Erickson, K. Escobar, A. Estes, M. Facchini, A. Falk, M. Farmer, B. Fearon, K. Ferguson, A. G. Fernandez, F. M. Fernandez, A. H. Ferrera, K. J. Ferris, B. Fett, A. Finck, J. Finley, L. E. Fischer, S. Flynn, F. F. Flöther, K. Fogel, J. P. Fouarge, G. Fraczak, P. Frank, D. Frank, J. C. Franquis, T. N. Friedhoff, D. Friedman, R. Frison, D. Fry, T. Fukui, B. Fuller, R. K. Fwamba, J. Gacon, M. Galante, J. M. Gambetta, C. A. Garcia, D. K. Garcia, J. M. Garcia, A. R. García, S. Garion, J. R. Garrison, B. P. Gaucher, A. Gebbie, J. Gerboth, S. Ghosh, F. Giancaspro, A. Giannetta, G. W. Gibson Jr., L. Gil, L. Gilgeous, G. Gionta, J. R. Glick, J. A. Glick, W. Glogowski, I. Goldman, J. T. Gomez, J. Gomez, S. de la Puente González, A. Gooding, R. T. Gordon, T. Greenfield, J. N. Greiner, L. Gruber, T. P. Gujarati, S. R. Gumede, K. Gunsky, H. Haas, R. Haight, S. A. Hall, I. Hamamura, J. B. Hannon, R. U. Haq, V. Harish, T. T. Harnack, S. Hart, G. S. Haspil, F. Haverkamp, V. Havlicek, C. Haymes, J. Haynes, A. He, M. B. Healy, R. Hebbar, J. Henspeter, F. J. C. Hernández, C. Herrick, J. Hertzberg, A. Hervier, H. Higgins, M. Hillenbrand, I. Hincks, C. W. P. Hoerst, K. L. Holland, B. Holt, M. Hopstaken, H. Horii, T. A. Horvath, R.

House, D. Hrtanek, E. Huang, J. Huang, J. R. Huffman, L. Hung, T. Huynh, N. Ibrahim, T. Imamichi, B. Ingmanson, K. Inoue, J. Irish, S. Itoh, T. Itoko, K. Jadin, H. Jagannathan, K. Jambunathan, J. A. Janechek, A. Javadi-Abhari, M. J. Jeanson, T. A. Jennings, O. Jinka, T. Jochym-O'Connor, B. R. Johnson, C. Johnson, S. Johnston, M. R. Jokar, B. A. Jones, E. A. Joseph, P. Jurcevic, H. Kang, M. Karakas, A. Kariye, S. Kashiwa, Z. Kauffman, D. Kaulen, J. Kaus, K. Kawase, G. A. Keefe, P. Keefer, N. Keenan, J. Kern, A. Khokhar, H. Kim, S. Kim, T. Kiraly, S. Kirschner, S. Kister, K. Knodel, W. Kong, B. Kreeger, M. Kriegshauser, K. Krsulich, V. Kumar, M. Kumph, C. Kurotori, D. Kwee, M. LaDue, R. Lallement, W. F. Landers, O. Lanes, G. J. Lapeyre, C. Larsen, O. Lasisi, I. Lauer, D. Layden, S. LeClerc, K. Lee, C. Lee, S. Lekuch, R. Letzter, M. Li, I. Liepuoniute, P. Lindner, T. Lindquist, G. Lindsell, E. G. Liniger, V. Lipinska, C. B. Lirakis, J. Lishman, D. Lokken-Toyli, R. Lolowang, M. Lu, W. Lu, S. Lukashov, A. Lukin, J. Luna, Y. Luo, K. Main, R. Majumdar, M. Malekakhlagh, R. Malik, H. Mallela, H. Mamin, C. Mancini, P. Mann, C. Marroquin, Y. Martin, I. D. Martinez, J. J. Martinez, Y. G. Martín-Mantero, N. A. Masluk, M. Mastro, K. Masuda, A. Matsuo, T. Maurer, D. McClure, M. McDonald, S. McJunkin, D. McKay, P. McKeone, K. McLaughlin, B. McPeek, D. Meirom, F. J. Mertil, R. Mertz, S. S. Meshram, A. Miessen, D. Millar, J. Milloz, Z. Minev, A. Mitchell, B. Mitchell, A. Mitra, T. Mittal, S. Miyabe, H. Mohammadbagherpoor, T. Mori, J. G. Mosquera, S. Mostame, M. A. Mueed, S. Mukesh, T. Mulder, Y. Muranaka, W. Murphy, C. Murray, J. P. Mzila-O'Connor, T. Müller, J. Nah, V. K. Naik, K. Najafi, D. Nakano, Y. Nam, P. Nation, J. S. Newbury, C. Nirkhe, S. Nur, P. O'Brien, K. O'Connell, L. E. Ocola, R. Ohira, M. Ohtani, T. Onodera, E. Opata, J. Orcutt, L. Oscarlece, R. Otaolea, I. Othmani, J. A. Ott, G. I. Ovejero, R. Padbury, E. B. H. Padilla, S. Panda, J. M. Papalia, R. Paranjape, P. Parazzoli, J. E. Park, B. Parney, R. T. Paske, D. Passarella, J. Patel, G. F. Paulik, E. Pednault, T. Pellegrini, B. Peropadre, K. S. Petrarca, S. Peyer, A. A. de la Peña, S. M. Pfaendler, M. Pfaffhauser, A. Phan, T. Phung, K. Pizzolato, B. Pokharel, T. Poole, K. Pooley, A. Popescu, C. Porter, A. Prabhakar, J. A. Prentice, E. Prew, E. Pritchett, M. Proissl, R. de Putter, D. Puzzuoli, A. Pyzyna, B. Quanz, E. Quigley, V. Radescu, J. J. Raftery, S. Raghunathan, A. Ragupathi, S. Rajalingam, P. Rall, J. Ralph, I. R. Ramallal, D. Ramirez, A. Rao, V. Rastunkov, A. Ray, S. Rayyan, C. T. Rettner, P. D. Reyes, C. Richard, D. Ristè, E. Rivera, P. Rivero, M. Rizzolo, J. M. Roberts, N. Robertson, N. Robinson-Duncan, K. Rodbell, D. J.

Rodriguez, D. M. Rodriguez, V. Rodriguez-Toro, A. Rojas, A. Rose, K. Ross, M. Rossmannek, T. Roth, M. B. Rothwell, J. R. Rozen, J. J. Ruedinger, D. Rugar, J. A. Russo, L. Sabaloz, H. Saenz, M. Sagianis, M. Sandberg, S. Sanichar, A. Sasaki, R. Sathananthan, A. Sauerland, H. Saunders, F. Scafirimuto, C. Scerbo, P. Schardt, M. J. Scheckel, L. Schleeper, R. Schmidt, J. M. Schmoll, R. Schoenberger, T. L. Scholten, K. Schoneck, J. Schuhmacher, A. Seif, M. Selvanayagam, H. H. R. Sepulveda, I. F. Sertage, P. Shah, W. Shanks, D. Shao, V. Sharma, D. Sharma, O. Shehab, R. M. Shelby, S. Sheldon, J. Sherman, N. Shimada, O. Shtanko, S. A. Shurson, V. Siddhu, W. Simmons, N. Simon, E. N. Sirong, J. W. Sleight, J. Smith, J. A. Smolin, B. Snell, D. Soederstedt, J. G. Sogo, J. Speidell, W. Spratt, S. Spring, S. Srinivasan, A. A. Stabile, K. Stawiasz, M. Steffen, J. Stehlik, R. Steiner, D. Stephenson, C. Stuckey, L. V. Subramaniam, J. Summerour, Y. Sun, N. Sundaresan, K. Sung, J. Suttle, T. Syed, C. Tabachnick, F. Tacchino, M. Takeori, M. Takita, A. Talle, C. Tan, E. P. Tapia, J. Tarm, I. Tavernelli, N. Taylor, J. Tersoff, S. Thomas, T. Thorbeck, S. Thoss, K. Tian, A. Tiano, K. Tien, J. Timmerwilke, T. Timpane, M. Tokunari, R. Tolentino, M. Tolunay, G. Toth, G. Totir, J. Totte, K. D. Tran, M. Tran, M. Treinish, B. D. Trimm, C. Tuma, M. Turchetti, L. Turlapati, C. S. Tyberg, P. Tysowski, G. Ulsh, D. Underwood, F. Valenti, D. G. Valiñas, M. VanDieren, F. Varchon, R. Vaught, A. C. Vazquez, A. Vilches, V. Villar, D. Vinayagamurthy, M. H. Vishkasougheh, S. K. L. P. Vishwakarma, C. C. Voss, C. Vu, E. Vásquez, A. Wack, M. Walther, C. Wang, D. Wang, F. Wang, M. Ware, J. Watrous, B. Webb, P. Wendling, B. Werneke, N. White, S. Willenborg, J. Williams, D. Wisnieff, S. Woerner, A. Wolseth, R. Woo, K. Wood, S. P. Wood, C. J. Wood, J. R. Wootton, J. L. Wright, B. Wymore, L. Xing, C. Xiong, T. Yamamoto, H. W. Yan, J. Yau, J. Yeh, K. Yoagaraj, T. Yoder, T. Younger, J. Yu, V. Yuen, D. Zajac, D. Zarsky, B. Zazueta, G. Zettles IV, E. J. Zhang, H. Zhang, R. Zhao, G. Zhu, S. Zhuk, S. Ziemak, C. Zoufal. M.Z. and S.A. were supported by the U.S. Department of Energy, Office of Science, Basic Energy Sciences, under Early Career Award No. DE-SC0022716. Y.W. is supported by the RIKEN iTHEMS fellowship. This work used the Anvil supercomputer at Purdue University through allocation PHY220016 from the Advanced Cyberinfrastructure Coordination Ecosystem: Services & Support (ACCESS) programme, which is supported by National Science Foundation grant nos. 2138259, 2138286, 2138307, 2137603 and 2138296. This research used resources of the National Energy

Research Scientific Computing Center, a DOE Office of Science User Facility supported by the Office of Science of the U.S. Department of Energy under contract no. DE-AC02-05CH11231 using NERSC award BES-ERCAP0024710. This research used the Lawrencium computational cluster resource provided by the IT Division at the Lawrence Berkeley National Laboratory (supported by the Director, Office of Science, Office of Basic Energy Sciences, of the U.S. Department of Energy under contract no. DE-AC02-05CH11231). We are indebted to Frank Pollmann and Sheng-Hsuan Lin for their contributions to the isoTNS code used in this work. S.A. thanks J. Hauschild for insightful conversations and support with the TeNPy library[33], which was used to implement the MPS simulations.

Author information

Author notes

1. These authors contributed equally: Youngseok Kim, Andrew Eddins

Authors and Affiliations

1. IBM Quantum, IBM Thomas J. Watson Research Center, Yorktown Heights, NY, USA

   Youngseok Kim, Ken Xuan Wei, Ewout van den Berg, Sami Rosenblatt, Hasan Nayfeh, Kristan Temme & Abhinav Kandala

2. IBM Quantum, IBM Research - Cambridge, Cambridge, MA, USA

   Andrew Eddins

3. Department of Physics, University of California, Berkeley, Berkeley, CA, USA

   Sajant Anand, Yantao Wu & Michael Zaletel

4. RIKEN iTHEMS, Wako, Japan

   Yantao Wu

5. Materials Sciences Division, Lawrence Berkeley National Laboratory, Berkeley, CA, USA

Michael Zaletel

Contributions

Y.K. and A.E. ran hardware and numerical experiments. S.A., Y.W. and M.Z. developed the tensor network methods and ran the classical simulations. Y.K., K.X.W. and A.K. developed the device-calibration protocol and calibrated the quantum processor with S.R. and H.N. E.v.d.B. and K.T. developed the error-mitigation protocol. Y.K., A.E., S.A., M.Z., K.T. and A.K. analysed the data. All authors contributed to the manuscript. K.T. and A.K. designed the project. A.K. managed the project.

Corresponding authors

Correspondence to Youngseok Kim, Andrew Eddins or Abhinav Kandala.

Ethics declarations

Competing interests

The authors declare no competing interests.

Peer review

Peer review information

Nature thanks Jonas Bylander, Zhenyu Cai, Balint Koczor, Göran Wendin and the other, anonymous, reviewer(s) for their contribution to the peer review of this work. Peer reviewer reports are available.

Additional information

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Supplementary information

Supplementary Information

Supplementary Sections 1–10, including Supplementary Figs. 1–13, Supplementary Table 1 and Supplementary References.

Peer Review File

Rights and permissions

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by/4.0/.

Reprints and Permissions

About this article

Cite this article

Kim, Y., Eddins, A., Anand, S. *et al.* Evidence for the utility of quantum computing before fault tolerance. *Nature* **618**, 500–505 (2023). https://doi.org/10.1038/s41586-023-06096-3

- Received: 24 February 2023
- Accepted: 18 April 2023
- Published: 14 June 2023
- Issue Date: 15 June 2023
- DOI: https://doi.org/10.1038/s41586-023-06096-3

Share this article

Anyone you share the following link with will be able to read this content:

Get shareable link

Sorry, a shareable link is not currently available for this article.

Copy to clipboard

Provided by the Springer Nature SharedIt content-sharing initiative

What IBM's result means for quantum computing

- Shamini Bundell
- Nick Petrić Howe

Nature Nature Podcast 14 Jun 2023

Quantum computer scales up by mitigating errors

- Göran Wendin
- Jonas Bylander

Nature News & Views 14 Jun 2023

This article was downloaded by **calibre** from https://www.nature.com/articles/s41586-023-06096-3

- Article
- Open Access
- Published: 14 June 2023

Hadaean to Palaeoarchaean stagnant-lid tectonics revealed by zircon magnetism

- John A. Tarduno ,
- Rory D. Cottrell,
- Richard K. Bono,
- Nicole Rayner,
- William J. Davis,
- Tinghong Zhou,
- Francis Nimmo,
- Axel Hofmann,
- Jaganmoy Jodder,
- Mauricio Ibañez-Mejia,
- Michael K. Watkeys,
- Hirokuni Oda &
- …
- Gautam Mitra

Show authors

Nature **volume 618**, pages 531–536 (2023)

- 2035 Accesses
- 233 Altmetric
- Metrics details

Abstract

Plate tectonics is a fundamental factor in the sustained habitability of Earth, but its time of onset is unknown, with ages ranging from the Hadaean to Proterozoic eons[1,2,3]. Plate motion is a key diagnostic to distinguish between plate and stagnant-lid tectonics, but palaeomagnetic tests have been thwarted because the planet's oldest extant rocks have been metamorphosed and/or deformed[4]. Herein, we report palaeointensity data from Hadaean-age to Mesoarchaean-age single detrital zircons bearing primary magnetite inclusions from the Barberton Greenstone Belt of South Africa[5]. These reveal a pattern of palaeointensities from the Eoarchaean (about 3.9 billion years ago (Ga)) to Mesoarchaean (about 3.3 Ga) eras that is nearly identical to that defined by primary magnetizations from the Jack Hills (JH; Western Australia)[6,7], further demonstrating the recording fidelity of select detrital zircons. Moreover, palaeofield values are nearly constant between about 3.9 Ga and about 3.4 Ga. This indicates unvarying latitudes, an observation distinct from plate tectonics of the past 600 million years (Myr) but predicted by stagnant-lid convection. If life originated by the Eoarchaean[8], and persisted to the occurrence of stromatolites half a billion years later[9], it did so when Earth was in a stagnant-lid regime, without plate-tectonics-driven geochemical cycling.

Main

Ideally, the presence or absence of a mobile lithosphere can be tested using palaeomagnetism, but even the best preserved oldest rocks on Earth have experienced metamorphism, and this places severe restrictions on the type of magnetic carriers that might retain primary signals. Single-crystal palaeointensity (SCP), whereby single mineral crystals that contain magnetic inclusions capable of recording the ancient field are studied, provides an approach to see through this metamorphism[10]. This method has been applied to document the past strength of the geomagnetic field using progressively older Archaean rocks[11,12,13]. The results are consistent with the few available palaeointensity studies of whole rocks of low metamorphic grade[14,15] and/or those having signals dominated by silicate-hosted magnetizations[16,17]. Extant rocks older than approximately 3.45 billion years

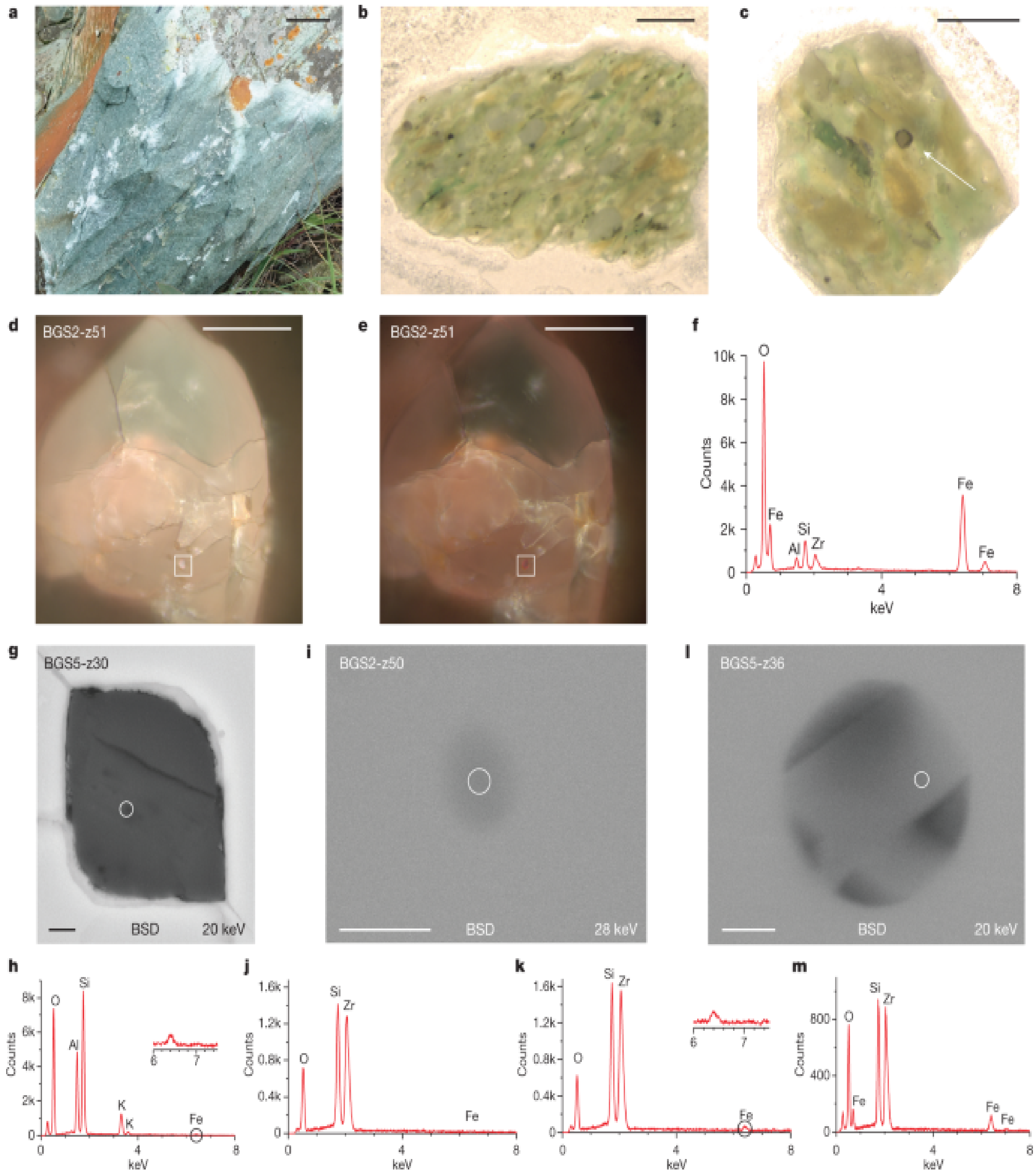
(Gyr) have generally been metamorphosed to amphibolite or higher grades, compromising any primary magnetic signals, but detrital single crystals found in younger sedimentary units have the potential to record even older geomagnetic fields[4]. The orientation of these detrital crystals at the time of their magnetization is unknown and, therefore, their magnetization direction does not constrain site latitude. The acquisition of palaeointensity data does not require knowledge of this orientation. SCP data can place bounds on motion because of the relationship between field strength and latitude, which is—in turn—a function of field morphology (Methods). Field-geometry constraints are unavailable before the Neoarchaean[18], but recent studies indicate the lack of an inner core for the timespan considered here[19,20]. Modelling of the geodynamo without an inner core suggests a dipole-dominated field[20].

The only known detrital crystals that can be accurately dated, and that are able to provide constraints on lithospheric mobility spanning the multi-hundred-million-year timescales that typify plate-tectonic cycles, are detrital zircons bearing primary magnetic inclusions[4]. However, the magnetic measurement of zircons and interrogations of their magnetizations to determine whether they preserve primary signals are formidable technical challenges[4,6,7]. The first SCP measurements of zircons were from the JH Discovery site of Western Australia[6]. Exhaustive tests led to selection of only approximately 2% of the zircons separated from their host rock, and these yielded palaeointensity data suggesting the presence of a geodynamo in the Hadaean eon, approximately 4.2 Ga (ref. [6]). This report was followed by a study showing how relatively young zircons successfully recorded the field[21] but also attempts to disprove the original findings[22,23]. Critiques have been systematically addressed[7,24,25,26,27,28] and, in so doing, further evidence on the preservation of primary magnetizations in zircons has been discovered. In particular, the presence of primary magnetite inclusions in select JH Discovery site zircons has been documented[7], consistent with magnetic unblocking temperature data presented in the first report on zircon magnetizations[6].

The distinctive time sequence defined by the JH palaeointensity values[6,7] is one of several lines of evidence indicating primary magnetizations rather than magnetic resetting (‘Evidence for primary magnetite inclusions and

magnetizations in JH zircons' section in Methods). However, before using the JH palaeointensity history to assess plate mobility, we have sought another record to further test its fidelity. Towards that goal, we sampled the Barberton Green Sandstone (BGS) of the Barberton Greenstone Belt, Kaapvaal Craton, South Africa (Fig. 1a–c), which has yielded detrital zircons ranging in age from 3.3 to 4.15 Gyr (refs. [5,29]). The BGS is less deformed and of lower metamorphic grade than the JH metaconglomerate zircon host rocks, with peak temperatures of ≤350 °C (ref. [30]). Palaeomagnetic conglomerate tests from the Barberton Greenstone Belt indicate the preservation of a high unblocking temperature primary magnetization[14,15].

Fig. 1: BGS, detrital zircons and Fe-bearing inclusions.



a, Field photo of BGS. Scale bar, 5 cm. **b**, Example section of BGS approximately 300 µm thick. Scale bar, 1 mm. **c**, Example of detrital zircon (arrow) in BGS matrix. Scale bar, 1 mm. **d**,**e**, Reflected-light images (100×) of zircon BGS2-z51, with the boxes highlighting Fe-oxide inclusion (magnetite) at 0° (**d**) and 90° (**e**) polarization. Scale bars, 50 µm. **f**, SEM EDS spectra of Fe inclusion highlighted in **d** and **e** (see Extended Data Fig. 2 for further SEM images and EDS analyses). **g**, SEM BSD image of silicate

inclusion in zircon BGS5-z30 with EDS analysis location identified by the circle (see Extended Data Fig. 1 for further SEM images and EDS spectra). Scale bar, 1 μm. **h**, EDS spectra of **g** with Fe signal highlighted. **i**, SEM BSD image of Fe-oxide inclusion (magnetite) in zircon BGS2-z50 with EDS spectra analysis location highlighted by the circle. Scale bar, 200 nm. **j**,**k**, EDS analyses of **i**. An Fe signal was not observed using a 20-keV beam (**j**) but was detected using a 28-keV beam (**k**), emphasizing that the inclusion is at depth (>4 μm). (Methods. Corresponding reflected-light images showing inclusion extinction are in Extended Data Fig. 2.) **l**, SEM BSD image of Fe-oxide particles in zircon BGS5-z36 with EDS location highlighted by the circle (further reflected-light and SEM images are shown in Extended Data Fig. 2). Scale bar, 200 nm. **m**, EDS spectra of **l**.

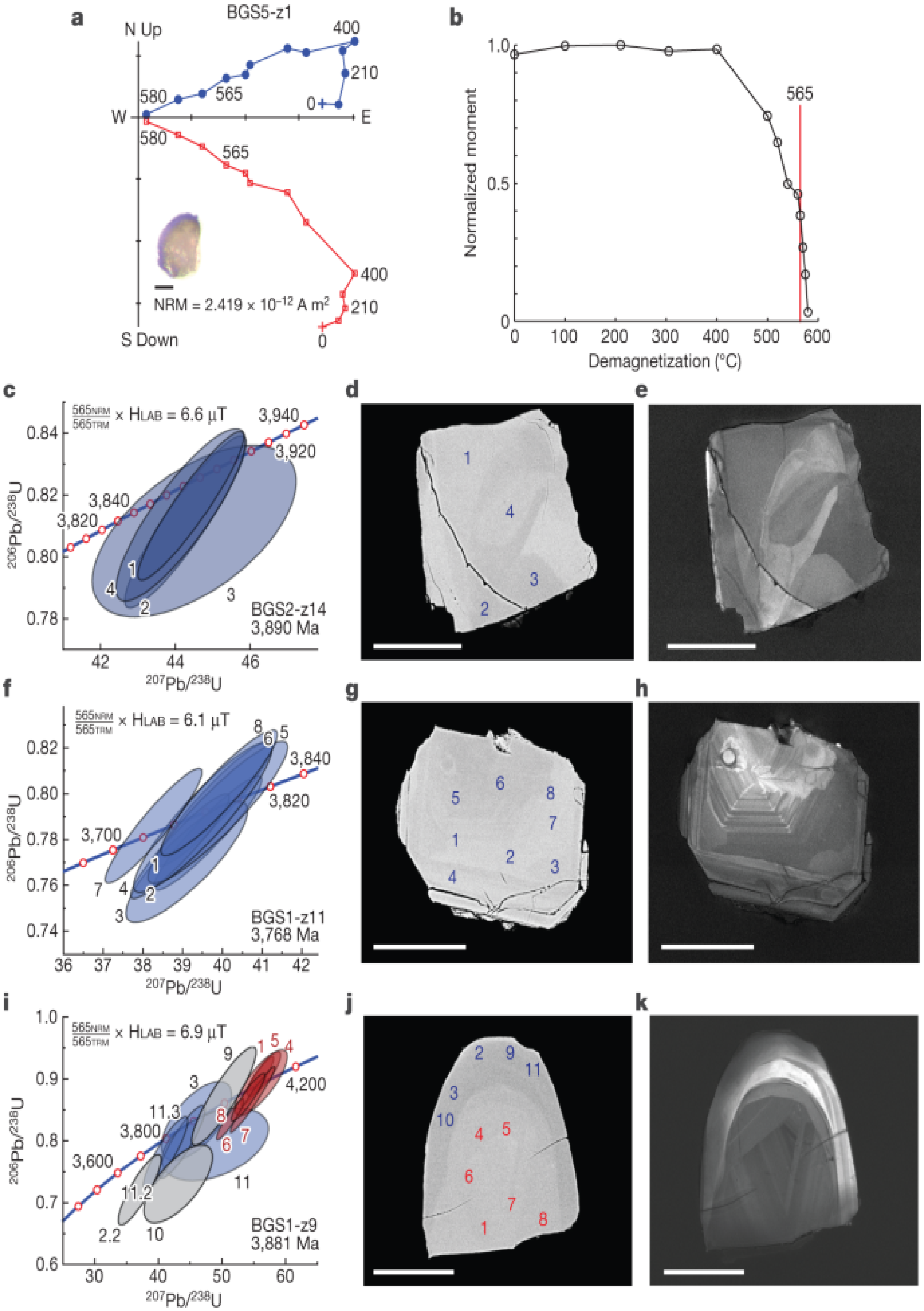
Source data

We separated zircons using a non-magnetic heavy-mineral technique (Methods). Reflected-light, scanning electron microscopy (SEM) and energy-dispersive X-ray spectroscopy (EDS; Methods) images show a diverse population of single-mineral inclusions and multiphase melt inclusions in the zircons, including apatite (Extended Data Fig. 1a–c), quartz (Extended Data Fig. 1d–g), feldspar (Fig. 1g,h and Extended Data Figs. 1h–n and 2a–f,h) and iron oxide (Fig. 1d–m and Extended Data Fig. 2f–j). Fractures record stress concentration around inclusions, providing further evidence that the inclusions are primary ('Microtectonic analyses of zircons' section in Methods and Extended Data Figs. 1 and 2). The iron-oxide inclusions show a variety of occurrences similar to those previously reported for JH zircons[6,7]. These include: (1) relatively large (>1 μm) particles (Fig. 1d,e and Extended Data Fig. 2f), (2) isolated iron-oxide particles at depth (Fig. 1i–k) and (3) Fe-oxide particles that crystallized in small (<1 μm) pockets (Fig. 1l,m). We also detected Fe that could be in the form of inclusions within larger silicates (Fig. 1g,h, Extended Data Fig. 1h–n and 'Reflected-light, SEM, palaeomagnetic and U-Pb analyses' section in Methods). For potential inclusions inside melt inclusions and/or silicates, we cannot assign an Fe-oxide type, but for others (that is, Fig. 1d–f,i–k and Extended Data Fig. 2g,i,j), reflected-light microscopy and EDS observations together indicate a magnetite composition ('Reflected-light, SEM, palaeomagnetic and U-Pb analyses' section in Methods). The Fe-oxide particles detected (for example, Fig. 1i,l) include those in the size/shape

range to have single-domain or stable single-vortex characteristics, with relaxation times of billions of years required to preserve primary signals.

For palaeointensity analysis, we adopt zircon selection criteria similar to those in our studies of JH zircons (Methods). Salient aspects of the analysis include the use of a CO_2 laser for demagnetization, allowing heating durations at least an order of magnitude less than those typical of conventional palaeomagnetic ovens, limiting laboratory alteration, and measurement with a WSGI ultrasensitive three-component direct-current superconducting quantum interference device (SQUID) magnetometer at the University of Rochester (Methods). The natural remanent magnetization (NRM) intensity of samples selected for analysis range from 9.4×10^{-13} A m^2 to 2.5×10^{-12} A m^2, representing approximately 13% of the more than 1,000 zircons separated (‘Reflected-light, SEM, palaeomagnetic and U-Pb analyses’ section in Methods). Demagnetization experiments yield results typical of magnetite, with complete unblocking at about 580 °C (Fig. 2a,b). On the basis of these results, we use the 565 °C palaeointensity method[6,7], developed to further limit laboratory alteration and to retrieve field-strength estimates to compare with the JH record (Methods).

Fig. 2: Palaeomagnetic thermal demagnetization, palaeointensity determinations and SHRIMP age data from individual BGS zircons.


a
BGS5-z1
N Up
W
E
S Down
400
210
580
565
0
NRM = 2.419 × 10−12 A m2
b
Normalized moment
Demagnetization (°C)
565
c
565NRM/565TRM × HLAB = 6.6 μT
BGS2-z14
3,890 Ma
206Pb/238U
207Pb/238U
d
e
f
565NRM/565TRM × HLAB = 6.1 μT
BGS1-z11
3,768 Ma
g
h
i
565NRM/565TRM × HLAB = 6.9 μT
BGS1-z9
3,881 Ma
j
k

a, Orthogonal vector plot of thermal demagnetization of unoriented zircon BGZ5-z1 (shown as inset). Temperatures shown are in °C. Red, vertical projection of the magnetization; blue, horizontal projection. Scale bar, 50 µm. **b**, NRM moment after thermal demagnetization (**a**) normalized to its undemagnetized value plotted versus demagnetization temperature. Value after heating at 565 °C highlighted (vertical red line). **c**, Concordia diagram showing SHRIMP geochronological analyses (uncertainty ellipses are 2σ) and palaeointensity value for zircon BGS2-z14. **d**, Corresponding backscattered scanning electron microscope image with analysed spots labelled. **e**, Corresponding cathodoluminescence image. **f**–**h**, Analyses as shown in **c**–**e** for zircon BGS1-z11. **i**–**k**, Analyses as shown in **c**–**e** for zircon BGS1-z9. Red, analyses from core; blue, analyses from rim; grey, excluded from calculation of mean age. Scale bars, 50 µm.

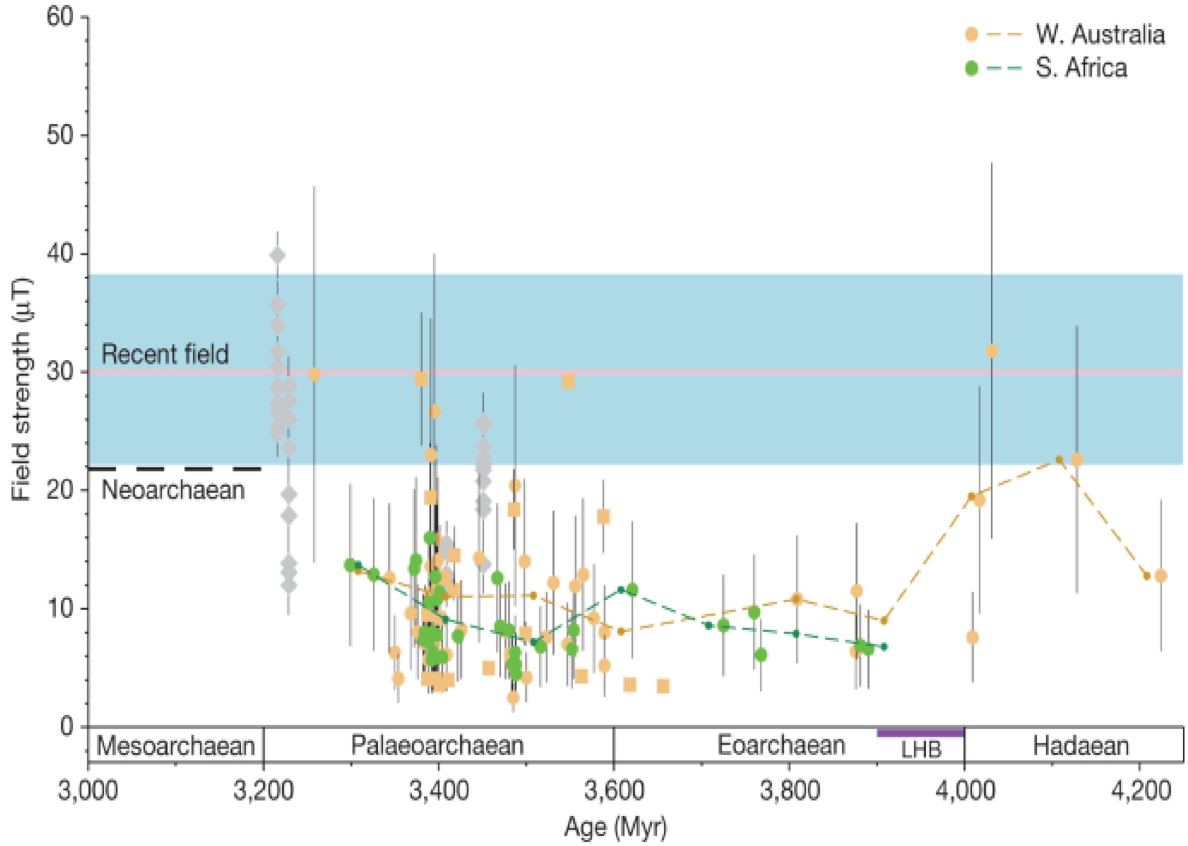
Source data

After magnetic analysis and the application of palaeointensity selection criteria (Methods), we analysed zircons using the Geological Survey of Canada SHRIMP II. Thirty-five zircons passed the selection criteria, representing approximately 3.5% of those separated (Methods), and their ^{207}Pb/^{206}Pb ages are considered further (Supplementary Table 1 and Supplementary Information). These yield Eoarchaean (Fig. 2) to Palaeoarchaean (Extended Data Fig. 3) concordant ages. Several zircons have older Hadaean-age cores (for example, Fig. 2i–k), extending back in time to 4.2 Ga (or older) (Extended Data Fig. 3j–l). However, these Hadaean zircons also have late Hadaean to early Eoarchaean, and sometimes younger, zircon rims (for example, Extended Data Fig. 3j–l). The late Hadaean to early Eoarchaean ages are intriguing because of their coincidence in time with the Late Heavy Bombardment[31]. In cases of evidence for growth of zircon rims, we assign the rim age with the time of magnetization.

Using this approach, we obtain a time sequence of palaeointensity values ranging from approximately 3.9 Ga to approximately 3.3 Ga (Fig. 3 and Extended Data Table 1). We find that we cannot reject the null hypothesis that these BGS data and those from the JH having ages between 3.9 and 3.4 Gyr sample the same underlying field distribution at the 95% confidence level using Welch's *t*, Kolmogorov–Smirnov (KS) and Mann–Whitney *U* tests ('Statistical analysis of BGS and JH zircon palaeointensity data' section

in Methods and Extended Data Table 2). For times younger than 3.4 Ga, however, Welch's *t*-test suggests that the BGS and JH palaeointensities begin to sample different distributions ('Statistical analysis of BGS and JH zircon palaeointensity data' section in Methods and Extended Data Fig. 4a). Although the data between 3.9 and 3.4 Ga are indistinguishable, the variance of the JH data is higher than that in the BGS data, which could reflect either a difference in recording fidelity and/or the time over which the field is recorded ('Statistical analysis of BGS and JH zircon palaeointensity data' section in Methods). Furthermore, the data are not evenly distributed in time, suggesting that only the broad character of the field over long timescales can be retrieved. Accordingly, we calculate 100-Myr time averages (Extended Data Table 3) to ensure that the time-averaged field needed to recover a dipole dynamo signal (palaeomagnetic dipole moments (PDMs), equations (1) and (2) in Methods) are represented. We find that the JH palaeointensity history based on the 100-Myr time window averaging (Extended Data Fig. 4b) predicts the BGS data (Extended Data Fig. 4c) over this same time range (KS test $P = 0.32$, above a significance level of 0.05; 'Statistical analysis of BGS and JH zircon palaeointensity data' section in Methods and Extended Data Fig. 4d–f). This similarity from widely separated sites with different Palaeoarchaean to recent geologic histories strongly support the interpretation that select JH and BGS zircons preserve a primary global signal of the geodynamo (also see the 'Reflected-light, SEM, palaeomagnetic and U-Pb analyses' section in Methods). But the unchanging 3.9–3.4-Ga palaeointensity values also indicate constant site latitudes.

Fig. 3: Palaeointensity history from BGS (South Africa) and JH (Western Australia) zircons.



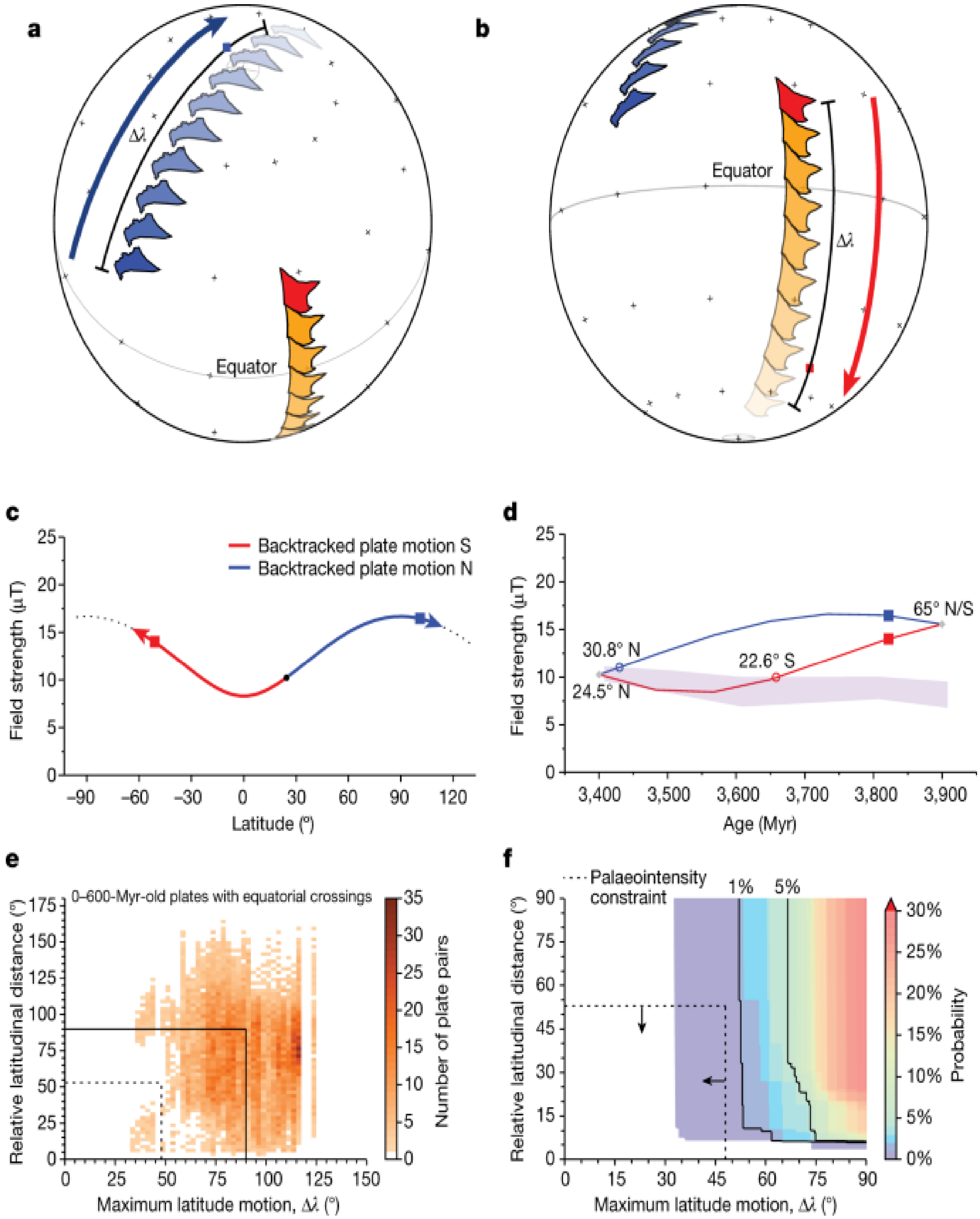
Zircon palaeointensity results: green circles, 565 °C palaeointensity determinations (this study); yellow from the JH (boxes, Thellier–Coe palaeointensity results; circles, 565 °C palaeointensity determinations)[6,7]. Green and yellow dashed lines: 100-Myr running average of zircon palaeointensity results from the BGS and JH, respectively. Other single-silicate palaeointensity results from extant igneous rocks shown as grey diamonds[12,13]. Recent field: pink solid line is mean and standard deviation (blue shaded region) from a bootstrap resampling of data[6] from the past 800 thousand years, set to the palaeolatitude of the Mesoarchaean data[12]. Neoarchaean field strength (dashed black line) based on mean of select time-averaged palaeointensity results[16,17]. LHB, Late Heavy Bombardment. All data are above threshold for geomagnetic field presence based on external field imparted by the solar wind[4]. Near-constant palaeointensity values between approximately 3.9 Ga and approximately 3.4 Ga indicate palaeolatitude stasis of the recording sites (see text).

Source data

We next test whether constant latitudes could arise from a modern style of plate tectonics by synthetically sampling continental areas and tracing their motion over the past 600 Myr using plate reconstructions[32] ('Plate-motion analysis' section in Methods and Extended Data Figs. 5 and 6). Geochemical data from JH zircons support an andesitic source, interpreted to record a modern arc-type setting[33], whereas those from the BGS support more heterogeneous crustal sources, including those distinct from that of the JH[29]. On this basis, we consider the case that the BGS and JH zircons were formed on at least two notional plates. However, no examples of two plates at constant latitudes could be found in the past 600 Myr of plate tectonic history.

We next investigate the resolution of the BGS/JH palaeointensity data in constraining latitudinal motion by considering hypothetical histories of plates moving back in time from 3.4 to 3.9 Ga ('Plate-motion analysis' section in Methods). We start at a palaeolatitude of approximately 24.5° constrained by the palaeomagnetism of extant Palaeoarchaean BGS rocks[13] (but arbitrarily set to the Northern Hemisphere) and allow the plates to traverse 90° of latitude, one backtracking south, the other north (Fig. 4a–d). The median maximum latitudinal travel for plates of the past 600 Myr is approximately 76°, different from the history inferred from the BGS/JH data (Fig. 4d).

Fig. 4: Hypothetical latitudinal motions typical of 0–600 Myr plate tectonics.



a, Two hypothetical continents, one backtracked north (blue) and the other south (red, orange), from a starting latitude of 24.5° corresponding to the 3.4–3.45-Ga palaeolatitude of the BGS[13] (hemisphere is unknown and arbitrarily set as N). Northward motion (blue arrow) and maximum

latitudinal motion (Δλ) highlighted. Blue square represents the median Δλ value observed for continental plates of the past 600 Myr. **b**, Same as **a** but the continent backtracked south is highlighted (in red and red arrow). Red square represents the median maximum latitudinal value observed for continental plates of the past 600 Myr. **c**, Dipole relationship between field intensity and latitude, set to the palaeolatitude of the BGS at 3.4–3.45 Ga (hemisphere is unknown and arbitrarily set as N), and the BGS/JH field strength value. Squares as in **a** and **b**. **d**, BGS and JH palaeointensity data (W. Australia + S. Africa) shown as the standard error of the 100-Myr bin mean (violet band) with predicted latitude history for a site moving northward (**a**) or southward (**b**). Squares as in **a** and **b**. Open circles are the values for which the backtracked palaeolatitude began to differ from the observations. **e**, 2D histogram of the relative latitudinal distance and maximum latitude motion (Δλ) characteristics of equatorial crossing plates of the past 600 Myr. Colour scale shows the number of unique plate pairs that exhibit motion as in **b**. Data are grouped into 2.5° bins. Dashed lines are the BGS/JH palaeointensity constraints (**d**; see 'Statistical analysis of BGS and JH zircon palaeointensity data' section in Methods). **f**, Expanded view of **e**, shown as the probability of sampling a pair of equatorial crossing plates with these relative latitudinal distance and Δλ characteristics. Dashed lines and arrows are the BGS/JH palaeointensity constraints.

Source data

The dipole relationship of field with latitude for a site originating at 24.5° creates an asymmetry in the resolving power of intensity data for a site backtracked to the north versus one backtracked to the south (Fig. 4c,d). Specifically, if we consider a site at 24.5° N, the BGS/JH intensity data are inconsistent with more than about 5° of latitudinal motion backtracked to the north (Fig. 4d). However, the same site backtracked to the south could have crossed the equator and have motion compatible with the data until it passed a latitude of 22.6° S, for a total latitudinal travel of approximately 47° (Fig. 4d). There is no example of a plate pair in the past 600 Myr that meets the requirements of the northward backtracked case. The probability of a plate pair meeting requirements of the southward backtracked possibility is <1% (Fig. 4e,f and 'Plate-motion analysis' section in Methods). Palaeolongitude is not directly constrained by our data, but our analyses show that plates with dominantly longitudinal motion are rare over the past 600 Myr ('Plate-

motion analysis' section in Methods). Moreover, because all plates show substantial latitudinal motion, our comparative tests inherently include these cases.

Discussion

The stark differences between the palaeolatitude history inferred for the BGS and JH sites between 3.9 and 3.4 Ga and Phanerozoic plate tectonics, highlighted by the comparisons above, provide evidence for a non-mobile lithosphere. Together with indications of the growth of continental crust from zircons[33,34], this indicates a regime of stagnant-lid tectonics[35] for Earth in Palaeoarchaean to Hadaean times, with plate tectonics, as defined by large-scale horizontal motions, possibly commencing after 3.4 Ga ('Prior geologic models and attendant implications' section in Methods). Processes to cool the mantle during the stagnant-lid regime should have included plumes and/or heat pipes[36]. We note that there is geological evidence in the Eoarchaean Itsaq Gneiss Complex of southwestern Greenland for crustal shortening[37]. However, our data suggest that this shortening and, more generally, the crustal recycling needed to generate continental crust[33,34] must have been spatially localized without extensive deep subduction creating highly mobile plates. Similarly, the lack of a mobile lithosphere indicates that potential subduction initiated by impacts[38] did not lead to a global system of plates and plate boundaries similar to that of today.

The Eoarchaean to Palaeoarchaean eras may be unique in Earth history because our data indicate that changes in geodynamo efficiency, which can otherwise dominate field strength, were minor ('Palaeointensity variations and dynamo efficiency' section in Methods). Thus, palaeointensity can provide first-order bounds on the locations of the first sites of crustal generation. In particular, we note that the palaeolatitude/latitude stasis, together with the latitudinal constraints at 3.4–3.45 Ga from extant rocks[13], indicate that the sites of crustal generation producing the BGS and JH zircons were at low latitudes from the Eoarchaean to Palaeoarchaean eras, and probably since the Hadaean eon.

The geochemical cycling provided by plate tectonics is recognized as being a key factor in sustaining habitability of Earth. However, it has been less

clear whether plate tectonics was required for the origin and early viability of life. Our new data suggest that, if life persisted throughout the Eoarchaean to Palaeoarchaean, it did so while Earth was in the stagnant-lid convective regime, demonstrating more broadly that a modern style of global plate tectonics is not a requirement for life on terrestrial planets during their first billion years. Moreover, the latitudinal history of the BGS and JH zircons indicates stability of the Earth relative to the spin axis, which is more likely in a planet without deep (>660 km) subduction[39]. The lack of large, rapid changes in environmental conditions induced by true polar wander likely fostered survival of nascent life on our planet.

Methods

Sample preparation and selection

Zircons are more dispersed in the BGS relative to zircons in the JH Discovery site samples studied by Tarduno et al.[6,7] and, therefore, separations were performed at the University of Rochester using non-magnetic gravimetric methods. Methylene iodide was the heavy liquid used and zircons were handpicked from the separate. Zircon selection follows the methods of Tarduno et al.[6,7], with some modifications. The BGS zircons isolated are generally slightly smaller than those studied from the JH Discovery site and, therefore, we focus on zircons larger than approximately 70 μm (versus 150 μm). The NRM intensity selection cutoff for analysis (applied to ensure that demagnetized values can be accurately measured) was slightly lower for the BGS zircons (approximately 9×10^{-13} A m^2 versus approximately 1×10^{-12} A m^2).

Viable zircons were collected at three sites (BGS1, BGS2 and BGS5). The separations yielded >1,000 zircons, but only about 500 met the size requirements. Of these, approximately 300 crystals could be isolated that were single zircons without attached non-zircon grains and that were clear of obvious surface alteration. One hundred and thirty zircons had NRM values $\geq 9 \times 10^{-13}$ A m^2. Sixty-five of these crystals without visible fractures or large inclusions (that might be multidomain oxides) were used for the 565 °C palaeointensity experiments. Seven zircons were rejected before

experiments because of non-reproducibility of initial NRM intensity measurements; the initial value was likely contaminated by a viscous remanent magnetization. A further five zircons failed because of a large intensity drop after the first thermal treatment, again suggesting a viscous remanent magnetization. Six zircons failed palaeointensity checks (they did not lose NRM on heating or did not gain thermoremanent magnetization (TRM) or failed the multidomain-tail check). Forty-seven zircons meeting palaeointensity criteria as established by Tarduno et al.[6] were subject to SEM and cathodoluminescence analyses and 12 were rejected because of metamictization or other evidence of internal grain disturbance. The 35 selected zircons represent a success rate of at least approximately 3.5% from those isolated from the methylene iodide heavy liquid separate. Although techniques differ, this success rate is higher than that of the study of JH zircons (<2%; Tarduno et al.[6]), consistent with a better preservation of the BGS zircons.

Optical and scanning electron microscopy

Optical microscopy studies used a Nikon Eclipse LV100POL microscope with both transmitted and reflected-light capabilities, a maximum of 1,000× magnification and a SPOT Insight 4-MP CCD colour digital camera assembly. SEM analyses were also conducted at the University of Rochester using a Zeiss Auriga SEM. For subsurface inclusions, we estimated depth using Electron Flight Simulator Version 3.1E software.

Palaeointensity experiments

Procedures follow those of Tarduno et al.[6,7] and are outlined here. All experiments were conducted in the shielded room (ambient field < 200 nT) of the Paleomagnetism Laboratory at the University of Rochester. The 565 °C palaeointensity method[6] was used to derive palaeofield strength estimates while limiting the effects of laboratory-induced alteration by reducing the number of heatings. A temperature of 565 °C is chosen for comparison with the JH palaeointensity dataset, which also used this temperature, and to represent blocking temperatures best reflecting single-domain-like magnetite inclusion carriers, while also retaining a measurable demagnetized value. The demonstration that 565 °C falls within the interval of sharp magnetic

unblocking (Fig. 2b) supports this choice. In the 565 °C palaeointensity approach, the NRM is first measured, and then the zircon was gradually heated in field-free space with a Synrad v20 CO_2 laser[12]. Specifically, the temperature was held constant for 1 min at 100-°C temperature steps to 500 °C. Next, the temperature was increased to 565 °C for an extra 1 min and then allowed to cool for 3–5 min. Thereafter, a second heating to 565 °C was conducted in the presence of a field. An applied field of 15 µT was used in all field-on treatments. A multidomain check was performed last by reheating the zircon sample to 565 °C in the absence of a field. We tabulate the calculated angle between the TRM vector (565_{on}-565_{off}; Extended Data Table 1) and the applied field. Differences from zero may record small error contributions related to the alignment of the laser beam and a zircon in these challenging experiments and/or an anisotropy of the collection of magnetic particles in a given zircon.

Relating palaeointensity to palaeolatitude

The geomagnetic field can be as described by the scalar potential ${\varPsi }_{m}(r, \theta, \phi, t)$:

$${\varPsi }_{m}(r,\theta ,\varphi ,t)=\frac{{r}_{{\rm{e}}}}{{\mu }_{0}}\mathop{\sum }\limits_{l=1}^{{\rm{\infty }}}\mathop{\sum }\limits_{m=0}^{l}{\left(\frac{{r}_{{\rm{e}}}}{r}\right)}^{l+1}{P}_{l}^{m}\cos \theta \,[{\,g}_{l}^{m}(t)\cos m\varphi +{h}_{l}^{m}(t)\sin m\varphi]$$
(1)

in which ${P}_{l}^{m}$ are partially normalized Schmidt functions, *l* and *m* are spherical harmonic degree and order, respectively, r_e is the radius of Earth and the Gauss coefficients ${g}_{l}^{m}(t)$ and ${h}_{l}^{m}(t)$ describe the spatially and time-varying fields. In the case of a time-averaged field needed to make conclusions on the geodynamo[10], the geocentric axial dipole (${g}_{1}^{0}$) is represented by a PDM and field strength *B* is related to palaeolatitude (λ) by:

$$B=\frac{{\mu }_{0}{\rm{P}}{\rm{D}}{\rm{M}}}{4{\rm{\pi }}{r}^{3}}\sqrt{1+3{\sin }^{2}\lambda }$$

(2)

Geochronology

Standard SHRIMP U-Pb instrumental setup[40]and procedure for U-Pb calibration by reference materials[41] were followed. Two 1-inch epoxy mounts (IP985 and IP987) were built containing the JH and BGS zircons as well as zircon reference materials z6266 with an accepted $^{206}Pb/^{238}U$ age of 559 Myr[41] and z1242 with an accepted $^{207}Pb/^{206}Pb$ age of 2,679.6 ± 0.2 Myr[42]. Polishing with diamond suspension exposed the internal structure of the zircon grains. After coating with 10 nA of gold, the grains were imaged in cathodoluminescence (CL) and backscattered electron detector (BSD) mode using a MIRA3 TESCAN field emission scanning electron microscope. In order to target pristine areas of zircon grains and avoid cracks and alteration, a ^{16}O- ion beam 13 μm in diameter and with an average beam current of 1 nA was used. The analytical runtable consisted of 11 masses including background with species of Hf, Yb, Zr, Pb, Th, U analysed over six scans. Data reduction used SQUID3 (ref. [43]) (note that the name of this software has no relation with SQUID magnetometers; SQUID3 refers to the name of the software and it is not an acronym). Steiger and Jäger[44] decay constants were used. Two analytical sessions were carried out on each of the two epoxy mounts. The 1σ external error for the $^{206}Pb/^{238}U$ calibration relative to reference material z6266 for sessions IP985_1 and IP985_2 was ±0.88% or ±1.01, respectively. The 1σ external error for the $^{206}Pb/^{238}U$ calibration for sessions IP987_2 and IP987_3 was ±0.75% or 1.07%, respectively. These errors are also specified in the footnote to the data table (Supplementary Table 1). The requirement for a correction due to instrumental mass fractionation of the Pb isotopes was assessed through replicate analyses of reference material z1242. The measured weighted mean $^{207}Pb/^{206}Pb$ age of those analyses are reported for each session in the footnotes of Supplementary Table 1. No fractionation correction was applied to the Pb- isotope data except for session IP987_2, in which a correction of −0.64% was applied. Concordia plots were generated, and weighted means calculated using Isoplot v. 4.15 (ref. [45]). The uncertainties for the weighted mean ages reported in the text and Supplementary Materials are at the 95% confidence level.

Evidence for primary magnetite inclusions and magnetizations in JH zircons

Palaeomagnetic, reflected-light microscope, electron microscope, microtectonic, geochemical and palaeointensity data indicate the presence of primary magnetite inclusions in select JH zircons and that these have primary magnetic signals[6,7]. Specifically, zircon microconglomerate tests and the identification of distinct secondary components of magnetization separate from the primary characteristic magnetization exclude magnetic resetting after incorporation of the zircons into the JH host metasediment (figure 2 of ref. [7]). ^{7}Li profiling data (for example, figure 6 of ref. [7]) argues against thermal resetting of Hadaean and Eoarchaean data, whereas reflected-light microscopy, microtectonic analysis, SEM/EDS, focused ion beam and NanoMOKE investigations document the presence of primary magnetite inclusions (figures 3–5 of ref. [7]). Pb-Pb screening of data, documented in refs. [6,7] (figure 7 of ref. [7]) argues against a magnetic resetting age older than the age of incorporation into the host sediment in the selected zircons. This is further enforced by the distinct change in palaeointensity data, with values from late Hadaean zircons being higher than those of Palaeoarchaean to Eoarchaean age, inconsistent with Palaeoarchaean magnetic resetting. Reference [6] notes that iron oxyhydroxides can be found in JH zircons that are likely related to weathering, but these cannot be the carrier of the magnetic remanence used for palaeointensity determination. Furthermore, an assemblage of magnetite inclusions is needed to account for the magnetizations observed from the zircon, satisfying Maxwell–Boltzmann statistical limits on magnetic recording[7]. Reference [7] also describes why remagnetization scenarios calling on the neoformation of magnetic grains represent particles too small or too few in number to record stable magnetizations. Reference [7] also notes that the association of magnetic particles with dislocations does not mean that the particles are secondary[22]; instead, this is expected given the deformation history of the JH conglomerate.

Although the zircon studied would not pass our selection criteria, atom probe tomography[46] data on a JH zircon illustrate differences in the high unblocking temperature magnetic primary remanences isolated in refs. [6,7]

and other potential Fe-bearing minerals. Reference [46] interprets quantum diamond microscope data from ref. [22] on a single zircon as indicating Fe-bearing zones and links those to the carriers of natural magnetic remanence. Atom probe tomography data from one zone within that zircon is further interpreted to contain Fe nanoclusters (which are far too small to carry remanent magnetizations) and a maximum age of approximately 1.4 Gyr is assigned; this is interpreted as a natural remanent magnetization age. A quantum diamond microscope does not have the sensitivity to measure the natural magnetic remanence of JH zircons. Reference [7] explains that the authors instead measured the magnetization of a laboratory-induced isothermal remanence magnetization (0.25 T, >4,000× present-day field; ref. [22]); this strong field can enhance the magnetization of Fe oxides/oxyhydroxides that do not contribute to the primary remanence. The zircon microconglomerate tests[6,7] discussed above supersede older tests and document that the high unblocking temperature magnetization in JH zircons must be older than the approximately 3-Gyr depositional age of the JH conglomerate. Therefore, the hypothetical remanent magnetization inferred in ref. [46] with an assigned age of approximately 1.4 Gyr cannot be the key high unblocking temperature component of JH magnetization[6,7].

Microtectonic analyses of zircons

Samples for this study were collected from the lower approximately 1 m of the approximately 3–5-m-thick BGS. Our samples appear distinct from those of the lowest 20 cm of the BGS, which show extensive shear zones in thin section[47]. In thin section, our BGS samples used for zircon separation have a greenschist-grade anastomosing foliation in the generally fine-grained matrix. The foliation is cut by several conjugate through-going fractures. The competency difference between the competent zircon grains and the weaker surrounding matrix could result in stress buildups at the grain boundaries that might generate cracks within the zircon grains. Below, reference is made to specific zircons studied by reflected-light and scanning electron microscopy.

SEM sample BGS2-z51. The fractures appear to be in three main systems (Extended Data Fig. 2a). The ‘vertical’ or N–S fractures are parallel to the *c* axis (crystallographic axis based on geometry). The ‘east-dipping’ and

‘west-dipping’ sets appear to be part of the same (112) crystallographic system as close-packed planes (for example, noted by Reddy et al.[48]). The melt inclusions (Extended Data Fig. 2b,f,d) are interesting because that in panel d does not have a fracture going through it, further indicating that the inclusions are not being formed by material diffusing in through the fractures. Instead, these patterns suggest that the inclusions are primary and many of the fractures are localized on the inclusions. The competency difference between the host crystal and the inclusions likely result in stress concentrations at the edges of the inclusions (depending on their composition). The fractures would tend to propagate along close-packed crystallographic planes (that is, the (112) planes).

Reflected-light sample BGS5-z30, Extended Data Fig. 1h,i. From the form of the crystal, the *c* axis is approximately vertical. There are two sets of conjugate fractures with the acute angle between them facing approximately E–W (they form the V pattern in the upper part of the grain), which would suggest an E–W compression direction when those fractures formed. The E–W fracture on either side of the ‘football-shaped’ inclusion probably formed during the same compression. The fractures coming off the tapered ends of the ‘football-shaped’ inclusion are likely guided by the inclusion itself. The tapered ends of an inclusion have the smallest radius of curvature and, therefore, generate the highest stress concentrations, causing the fracture to propagate out from the tapered tip. Those fractures do not appear to be guided by the crystal structure of the zircon—if the fractures had propagated farther, they might have reoriented to follow a crystallographic close-packed plane. There is a possible fracture within the inclusion (Extended Data Fig. 1j) that could reflect propagation from a stress concentration outside the inclusion.

Reflected-light sample BGS2-z39, Extended Data Fig. 1. Although the fractures look random at first glance (Extended Data Fig. 1d), a few fractures seem to be parallel to the *c* axis or 001 (along the length of the crystal), a set is perpendicular to the *c* axis (on the bottom of the crystal) and there is a pair of conjugate sets (perhaps the (112) set) indicating compression approximately in the E–W direction. The globular form of the quartz inclusion highlighted in Extended Data Fig. 1e,f suggests that it is a melt inclusion. The adjacent fracture does not pass through the inclusion. The fractures likely originate at the inclusion boundary as a result of stress

concentration created by the inclusion. The fracture below the inclusion is clearly following a crystallographic plane parallel to the *c* axis.

SEM sample BGS2-z36, Fig. 11 and Extended Data Fig. 2k–l. The inclusion is not connected to fractures in the 2D view available. Lattice diffusion (Nabarro–Herring creep) requires temperatures greater than half the melting temperature. In the case of zircon, that would be approximately 900 °C, unless the crystal structure was sufficiently disturbed that the lattice was more open. There is no evidence of metamictization that might signal such an open structure. The distance from the edge of the crystal to the inclusion (in 2D) is approximately 2 μm and it is unlikely that the inclusion and interior Fe particles could form by diffusion through the crystal lattice. Instead, it is likely that this is a melt inclusion formed during the original formation of the crystal (at temperatures >1,000 °C).

Reflected-light, SEM, palaeomagnetic and U-Pb analyses

Magnetic inclusions on the tens to hundreds of nanometres scale are well established within silicates (for example, Tarduno et al.[10]); trace Fe signals could reflect such inclusions within the zircons, but further work is needed to distinguish these from Fe intrinsic to the crystals in the case of feldspar. When viewed under reflected light, we note that isolated inclusions often show extinction after rotation from 0° to 90° polarization angles. For those shallow enough in the crystal to be sampled by SEM EDS analysis, the presence of Fe together with extinction provides evidence for magnetite[7].

We emphasize that the zircons in our study were selected using robust protocols for the isolation of primary magnetic recorders, established in the first palaeomagnetic study of zircons[6]. Our NRM and demagnetization data, together with our transmitted/reflected-light and SEM observations, show that a claim that BGS zircons “contain virtually no ferromagnetic minerals”[49] is incorrect.

To demonstrate the robustness of the isotopic system of individual zircon grains, replicate analyses were carried out wherever possible. Zircon grains with reproducible Pb-Pb ages are inferred to be closed systems that have not lost Pb due to diffusion, recrystallization or other thermal processes. When

the probability of fit of the weighted mean age is greater than 0.05, this age is considered to be the magnetization age. In cases in which zircon overgrowths are present, we assign the youngest robust age with the time of magnetization. A description of the results of each individual grain is available in the Supplementary Information. Overall, we note that our BGS zircons lack 3.4 rims. This is consistent with our interpretation that these zircons have not experienced a 3.4-Ga high-temperature event that might otherwise have affected the magnetic history they record.

Statistical analysis of BGS and JH zircon palaeointensity data

To compare the JH and BGS zircon palaeointensity data, we first examine the following hypotheses: H_0, the JH and BGS data sample the same underlying field distribution, designated as the null hypothesis; and H_1, the JH and BGS data sample different underlying field distributions, designated as the alternative hypothesis. We focus on three statistical tests: the Welch version of the Student *t*-test (Welch[50]), the two-sample KS test[51] and the Mann–Whitney *U* test[52]. The Student *t*-test differs from the KS and Mann–Whitney *U* tests in the assumption that each sample is normally distributed with equal variance. Given the small sample sizes for some comparisons and unequal variances in each population (see below), the two-tailed Welch's version of the *t*-test is appropriate[50], which compares each sample mean and variance to the *t* distribution. The non-parametric two-sample KS test compares two samples and measures the maximum distance between their empirical cumulative distributions (D_{KS}) and rejects the null hypothesis if p_{KS} exceeds some critical threshold. The Mann–Whitney *U* test is another non-parametric test that examines two samples (for example, *X* and *Y*), with the null hypothesis being that a random value from sample *X* has an equal probability of being greater or less than a random value from population *Y*; this is the expected outcome if the two samples share the same underlying distribution. All three tests compare the hypotheses H_0 and H_1 and return a *P* value. If the *P* value is greater than a defined significance threshold (α), typically $\alpha = 0.05$, then the H_0 hypothesis cannot be rejected; if the *P* value is less than the defined significance, then H_0 can be rejected at $(1 - \alpha)$ confidence in favour of H_1. For our hypotheses, if $P > 0.05$, then it is

reasonable to infer that the JH and BGS zircon palaeointensity distributions cannot be distinguished at the 95% confidence threshold.

We first compare the JH and BGS between 3.4 and 3.9 Ga and justify this choice of age range further below; we find that the Welch's *t*-test, non-parametric two-sample KS test and Mann–Whitney *U* test all yield *P* values > 0.05, indicating that we cannot reject the null hypothesis H_0 (Extended Data Table 2).

In Tarduno et al.[7], JH zircon palaeointensity data were averaged in a moving-window model using 100-Myr non-overlapping bins to estimate the PDM for each 100-Myr interval. Bin edges were defined by the distribution of available data in their set of JH zircon palaeointensities, starting at 3,258 Ma, 3,358 Ma up to 4,258 Ma. We follow this approach to define 100-Myr averaging bins with the BGS data, using identical bin definitions to allow for direct comparison between JH and BGS palaeointensity data. Finally, palaeointensities from both JH and BGS datasets are combined using the same bin definitions to produce 100-Myr averages used in Extended Data Fig. 4. Palaeointensity statistics for each 100-Myr bin are provided in Extended Data Table 3.

We next consider the 100-Myr age bins defined in the moving-window model separately (Fig. 3 and Extended Data Table 2). Only two intervals contain more than five observations (the minimum threshold for Welch's *t*-test) for both the JH and BGS datasets, centred on 3,408 Ma and 3,508 Ma. The Welch's *t*-test for each interval yields *P* values of 0.17 and 0.09, respectively, and thus again the null hypothesis H_0 that the JH and BGS data record the same field intensity cannot be rejected. Non-parametric tests yield similar results, with *P* values exceeding 0.05 for both intervals (Extended Data Table 2), supporting the interpretation of the Welch's *t*-test.

To define the longest age interval that JH and BGS data sample the same underlying field, a series of Welch's *t*-tests was conducted for data 3.9 Ga and younger. For each test, the age interval was successively shortened by adjusting the younger bound back in time, with the resulting *P* value recorded (Extended Data Fig. 4a). We find that, for ages older than 3.4 Ga, the *P* value is high and supports the null hypothesis. We also note that, for age bounds less than 3.4 Ga, the Welch's *t*-test *P* value drops below 0.05,

resulting in the rejection of the null hypothesis H_0 in favour of H_1 (Extended Data Fig. 4a). This suggests that the JH and BGS data begin to sample different underlying field distributions when the comparison interval extends into younger times.

We next consider whether the JH palaeointensity history predicts that observed from the BGS between 3.9 and 3.4 Ga. We examine the residuals between the 100-Myr smoothing window model for the JH palaeointensity and the data used to define it (that is, the palaeointensity data from the JH zircons; Extended Data Fig. 4b,d). We do the same for the BGS palaeointensity dataset (that is, the BGS palaeointensity data versus the JH model; Extended Data Fig. 4c,e). Given the two distributions, we apply a two-sample KS test (Extended Data Fig. 4f) and assess whether the new data residuals look like the data that were input into the JH smoothed window average.

The null hypothesis is that the two samples (that is, residual populations) are drawn from the same distribution. In the case in which the new data are well predicted by the model, the KS test should return a P value > 0.05 and, if the new data are poorly predicted, the P value is <0.05 (that is, in this case, the hypothesis that the residuals are similar can be rejected). When applied to the new BGS, the P value is 0.32, which supports the hypothesis that the JH model can predict the Barberton zircon results. Although the P value is above the 95% threshold, the distributions nevertheless merit some discussion. The JH palaeointensities are clearly more scattered than the BGS results. This could signal one, or both, of the following:(1) the BGS zircons are, in general, better preserved, having less amorphous iron oxides in cracks than the JH and, therefore, they might be better recorders; (2) the JH might record the field on shorter timescales than the BGS, averaging higher-frequency secular variations of the past geomagnetic field. Palaeointensities from the SCP analyses of Nondweni dacite samples are within the BGS values of the same age, but SCP of Barberton dacites at 3.45 Ga are higher. Both the Nondweni and Barberton SCP values are plotted as individual results from two relatively shallow intrusions, and these might be expected to sample higher-frequency variations of the geomagnetic field than the zircons.

Plate-motion analysis

Here we consider the probability of observing two sites, located on different plates, experiencing both little to no latitudinal motion as well as little to no relative latitudinal distance separation during an interval of 600 Myr. To construct this test, the plate-reconstruction model of Merdith et al.[53], a continuous full-plate reconstruction model spanning the past 1 Gyr, is used. A set of sites are defined for the present day and the site palaeolocations are reconstructed back from the present day to 600 Ma using GPlates[54]. The distribution of palaeolocations is described and comparisons between sites located on different plates is made. From this empirical dataset, a set of statistical tests is constructed.

The sampling grid is defined using a Fibonacci spiral[55], which yields a uniform distribution of locations distributed globally with a median separation of approximately 6°. From this grid, 1,000 sites are assigned to plates using a built-in GPlates function. Sites that do not fall within the boundaries of a plate are removed before the analysis. Using the plate-motion model of Merdith et al.[53], site palaeolocations are reconstructed in 1-Myr steps from the present day to 600 Ma. From the initial set of sites, only those assigned to plates that existed 600 Ma are preserved, resulting in a set of 228 locations (Extended Data Fig. 5a).

We model a plate-motion path for each site from the GPlates reconstruction. Using the plate-motion path, the absolute maximum latitudinal distance travelled, $\Delta\lambda$ can be determined as:

$$\Delta \lambda =\max \left(| {\lambda }_{0}-{\lambda }_{{P}_{t}}| \right)$$
(3)

in which λ_0 is the present-day site latitude, $({\lambda }_{{P}_{t}})$ is the palaeolatitude at time *t* and ‘max’ is the maximum distance for the set considered. $\Delta\lambda$ distances range from 33° to 127°, with a median $\Delta\lambda$ distance of 76° (Extended Data Fig. 6b).

From the distribution of $\Delta\lambda$ for each of the 228 sites located on 66 unique 'plates' as designated by GPlates, the general trend of latitudinal motion can be described (Extended Data Fig. 5c). Broadly speaking, there is a weak positive correlation between the magnitude of $\Delta\lambda$ and the age at which the maximum latitudinal distance is observed (Pearson's correlation coefficient, $r = 0.49$);, however, sites with $\Delta\lambda$ falling in the lower 5% of the distribution ($\Delta\lambda \leq 40°$, $n = 12/228$) have ages ranging from 150 to 480 Ma. We note that $\Delta\lambda$ estimates can be biased downward by as much as approximately 10° when sites are downsampled at 100-Myr versus 1-Myr intervals.

With these general trends in mind, we next consider whether any sites experience little to no latitudinal motion. No sites could be identified with near-zero latitudinal motion irrespective of the sampling level (Extended Data Fig. 6a,b). If we arbitrarily define 'little motion' as ≤30° of latitudinal motion, no sites at 1-Myr or 20-Myr sampling meet the criterion, whereas only a single site (located on the Paraná/Pampia plate) met the criterion with 100-Myr downsampling. If the definition of 'little' motion is further relaxed to a threshold of ≤35° and the sampling interval to 100 Myr, 11 of 228 sites meet the criterion. The 11 sites showing ≤35° motion are observed on seven plates. We note that we only consider the maximum amount of latitudinal motion in this analysis; if both northern and southern latitudinal motion are considered separately instead, the lower 5th percentile of the revised distribution indicates a minimum total latitudinal motion of approximately 50–60°.

Given the rarity, or absence, of plates showing no motion over 600 Myr, the likelihood of observing two sites on separate plates with little to no latitudinal motion is very low. We next consider the bounds of $\Delta\lambda$ and relative latitudinal distance for plate pairs provided by the palaeointensity data (Fig. 4). The case of a plate backtracking to the north is automatically recognized as being inconsistent with plate motion over the past 600 Myr by the single-site $\Delta\lambda$ distributions (Extended Data Fig. 6a,b). For a plate backtracking to the south, the $\Delta\lambda$ bound is 48° (see main text). The bound on relative distance between the plate pair is the combination of this backtracked southward motion for one plate and the maximum backtracked northward motion of another plate consistent with the palaeointensity data (5°), for a total of 53°.

To determine the probability of sampling sites from two separate plates with these $\Delta\lambda$ and relative distance characteristics, a hypergeometric distribution is appropriate. A hypergeometric distribution models a sampling without replacement, in contrast to the more common binomial distribution, which samples trials with replacement[56]. Sampling without replacement is appropriate for this scenario because the same site (plate) cannot be selected twice. The probability mass function ($p(k)$) for a hypergeometric function is defined as:

$$p(k)=\frac{\left(\begin{array}{c}K\\ k\end{array}\right)\left(\begin{array}{c}N-K\\ n-k\end{array}\right)}{\left(\begin{array}{c}N\\ n\end{array}\right)}$$
(4)

in which $N = 66$ is the number of plates, K is the number of plates in the 600-Myr dataset known to record latitudinal motion and relative distance less than or equal to a specified threshold, $n = 2$ represents the number of unique plates selected in a single random sampling and k is the number of plates in that sampling showing limited latitudinal motion. The resulting probability of identifying a single pair of plates ($k = 2$) that both show limited latitudinal motion, here defined as $\Delta\lambda \leq 48°$ and relative latitudinal distance ≤53° at 100-Myr downsampling, is <1% (Fig. 4f).

Finally, we note that palaeolongitudinal trends characteristic of plate tectonics can also be inferred from plate-motion models spanning the past 600 Myr. To avoid biasing of longitudinal motion owing to geometric effects (that is, less area at higher latitudes), we normalize longitudinal motion by analysing the fraction of maximum angular travel from the initial position that is not explained by latitudinal change. We note that plates having dominant longitudinal motion (for example, >50%) are rare (about 5%). In these cases, there is at least 40° of latitudinal motion.

Prior geologic models and attendant implications

A 'Vaalbara' supercraton involving the Pilbara and Kaapvaal cratons between about 3.1 and 2.7 Ga has been proposed by some authors[57,58] but challenged by others[59]. Arguably the strongest geologic correlations are

between the approximately 2.7–2.8-Ga volcanism on these cratons (Ventersdorp Supergroup of the Kaapvaal Craton and Fortescue Group of the Pilbara Craton). Palaeomagnetic data for the Kaapvaal Craton, applying a tilting interpretation[59], imply emplacement at a very high latitude (75.4°). There is an approximately 700-Myr-long age gap between this volcanic association and the time when the BGS and JH zircon records start to diverge, which might hint at a start of horizontal motions characteristic of plate tectonics. This duration is 100 Myr longer than that over which we used to assess the latitudinal-motion characteristics of plate tectonics. Nevertheless, our analysis can be used as a conservative measure to assess whether plate-motion rates typical of plate tectonics are compatible with a high-latitude Kaapvaal Craton in the Neoarchaean and a low-latitude position in the Palaeoarchaean. Given that the median maximum latitudinal displacement from 0–600 Ma is 76°, it is clear that plate-motion rates are compatible. Assuming a location of the Kaapvaal craton solely in one (N or S) hemisphere yields an average latitudinal component of motion of 8 mm $year^{-1}$. Assuming one equatorial crossing results in an average latitudinal component of motion of 16 mm $year^{-1}$.

Palaeointensity variations and dynamo efficiency

Palaeomagnetic variations, principally the occurrence of reversing and seemingly non-reversing periods (superchrons), but also inferences on secular variation, have long been used to argue for changes in core–mantle boundary structure influencing the efficiency of the geodynamo[60,61]. Palaeointensity variations of superchron versus reversing intervals[62,63,64,65] have also been interpreted as reflecting these variations, and this is supported by some numerical geodynamo models[66]. More recently, it has been suggested that such changes may extend to the Devonian[67] and/or might have started in the Ediacaran Period[19]. If palaeointensity variations similar to those observed in the Phanerozoic had been present in our 3.9–3.4-Ga data, we might not have been able to separate changes in dynamo efficiency from palaeolatitude changes. But because the palaeointensity record is constant, we infer that palaeolatitudes were constant and mantle processes were not creating the latitudinal core–mantle boundary heat-flux patterns necessary to greatly affect geodynamo efficiency.

Data availability

Data presented here are available in the Earthref (MagIC) database (earthref.org/MagIC/19526; https://doi.org/10.7288/V4/MAGIC/19526). Source data are provided with this paper.

References

1. Hawkesworth, C. J., Cawood, P. A. & Dhuime, B. The evolution of the continental crust and the onset of plate tectonics. *Front. Earth Sci.* **8**, 326 (2020).

2. Shirey, S. B. & Richardson, S. H. Start of the Wilson cycle at 3 Ga shown by diamonds from subcontinental mantle. *Science* **333**, 434–436 (2011).

3. Stern, R. J. Evidence from ophiolites, blueschists, and ultrahigh-pressure metamorphic terranes that the modern episode of subduction tectonics began in Neoproterozoic time. *Geology* **33**, 557–560 (2005).

4. Tarduno, J. A., Blackman, E. G. & Mamajek, E. E. Detecting the oldest geodynamo and attendant shielding from the solar wind: implications for habitability. *Phys. Earth Planet. Inter.* **233**, 68–87 (2014).

5. Byerly, B. L. et al. Hadean zircon from a 3.3 Ga sandstone, Barberton greenstone belt, South Africa. *Geology* **46**, 967–970 (2018).

6. Tarduno, J. A., Cottrell, R. D., Davis, W. J., Nimmo, F. & Bono, R. K. A Hadean to Paleoarchean geodynamo recorded by single zircon crystals. *Science* **349**, 521–524 (2015).

7. Tarduno, J. A. et al. Paleomagnetism indicates that primary magnetite in zircon records a strong Hadean geodynamo. *Proc. Natl Acad. Sci. USA* **117**, 2309–2318 (2020).

8. Rosing, M. T. ^{13}C-depleted carbon microparticles in >3700-Ma sea-floor sedimentary rocks from West Greenland. *Science* **283**, 674–676 (1999).

9. Walter, M. R., Buick, R. & Dunlop, J. S. R. Stromatolites 3,400–3,500 Myr old from the North Pole area, Western Australia. *Nature* **284**, 443–445 (1980).

10. Tarduno, J. A., Cottrell, R. D. & Smirnov, A. V. The paleomagnetism of single silicate crystals: recording geomagnetic field strength during mixed polarity intervals, superchrons, and inner core growth. *Rev. Geophys.* **44**, RG1002 (2006).

11. Smirnov, A. V., Tarduno, J. A. & Pisakin, B. N. Paleointensity of the early geodynamo (2.45 Ga) as recorded in Karelia: a single-crystal approach. *Geology* **31**, 415–418 (2003).

12. Tarduno, J. A., Cottrell, R. D., Watkeys, M. K. & Bauch, D. Geomagnetic field strength 3.2 billion years ago recorded by single silicate crystals. *Nature* **446**, 657–660 (2007).

13. Tarduno, J. A. et al. Geodynamo, solar wind and magnetopause 3.45 billion years ago. *Science* **327**, 1238–1240 (2010).

14. Usui, Y., Tarduno, J. A., Watkeys, M. K., Hofmann, A. & Cottrell, R. D. Evidence for a 3.45-billion-year-old magnetic remanence: hints of an ancient geodynamo from conglomerates of South Africa. *Geochem. Geophys. Geosyst.* **10**, Q09Z07 (2009).

15. Biggin, A. J. et al. Palaeomagnetism of Archaean rocks of the Onverwacht Group, Barberton Greenstone Belt (southern Africa): evidence for a stable and potentially reversing geomagnetic field at ca. 3.5 Ga. *Earth Planet. Sci. Lett.* **302**, 314–328 (2011).

16. Selkin, P. A., Gee, J. S., Meurer, W. P. & Hemming, S. R. Paleointensity record from the 2.7 Ga Stillwater Complex, Montana. *Geochem. Geophys. Geosyst.* **9**, Q12023 (2008).

17. Muxworthy, A. R., Evans, M. E., Scourfield, S. J. & King, J. G. Paleointensity results from the late-Archaean Modipe Gabbro of Botswana. *Geochem. Geophys. Geosyst.* **14**, 2198–2205 (2013).

18. Smirnov, A. V. & Tarduno, J. A. Secular variation of the Late Archean–Early Proterozoic geodynamo. *Geophys. Res. Lett.* **31**, L16607 (2004).

19. Bono, R. K., Tarduno, J. A., Nimmo, F. & Cottrell, R. D. Young inner core inferred from Ediacaran ultra-low geomagnetic field intensity. *Nat. Geosci.* **12**, 143–147 (2019).

20. Driscoll, P. E. Simulating 2 Ga of geodynamo history. *Geophys. Res. Lett.* **43**, 2016GL068858 (2016).

21. Sato, M. Rock-magnetic properties of single zircon crystals sampled from the Tanzawa tonalitic pluton, central Japan. *Earth Planets Space* **67**, 150 (2015).

22. Tang, F. et al. Secondary magnetite in ancient zircon precludes analysis of a Hadean geodynamo. *Proc. Natl Acad. Sci. USA* **116**, 407–412 (2018).

23. Borlina, C. S. et al. Reevaluating the evidence for a Hadean-Eoarchean dynamo. *Sci. Adv.* **6**, eaav9634 (2020).

24. Bono, R. K., Tarduno, J. A. & Cottrell, R. D. Comment on: Pervasive remagnetization of detrital zircon host rocks in the Jack Hills, Western Australia and implications for records of the early dynamo, by Weiss et al. (2015). *Earth Planet. Sci. Lett.* **450**, 406–408 (2016).

25. Cottrell, R. D., Tarduno, J. A., Bono, R. K., Dare, M. S. & Mitra, G. The inverse microconglomerate test: further evidence for the preservation of Hadean magnetizations in metasediments of the Jack Hills, Western Australia. *Geophys. Res. Lett.* **43**, 4215–4220 (2016).

26. Dare, M. S. et al. Detrital magnetite and chromite in Jack Hills quartzite cobbles: further evidence for the preservation of primary magnetizations and new insights into sediment provenance. *Earth Planet. Sci. Lett.* **451**, 298–314 (2016).

27. Bono, R. K., Tarduno, J. A., Dare, M. S., Mitra, G. & Cottrell, R. D. Cluster analysis on a sphere: application to magnetizations from metasediments of the Jack Hills, Western Australia. *Earth Planet. Sci. Lett.* **484**, 67–80 (2018).

28. Bono, R. K., Tarduno, J. A. & Cottrell, R. D. Primary pseudo-single and single-domain magnetite inclusions in quartzite cobbles of the Jack Hills (Western Australia): implications for the Hadean geodynamo. *Geophys. J. Int.* **216**, 598–608 (2019).

29. Drabon, N. et al. Heterogeneous Hadean crust with ambient mantle affinity recorded in detrital zircons of the Green Sandstone Bed, South Africa. *Proc. Natl Acad. Sci. USA* **118**, e2004370118 (2021).

30. Tice, M. M., Bostick, B. C. & Lowe, D. R. Thermal history of the 3.5–3.2 Ga Onverwacht and Fig Tree Groups, Barberton greenstone belt, South Africa, inferred by Raman microspectroscopy of carbonaceous material. *Geology* **32**, 37–40 (2004).

31. de Niem, D., Kührt, E., Morbidelli, A. & Motschmann, U. Atmospheric erosion and replenishment induced by impacts upon the Earth and Mars during a heavy bombardment. *Icarus* **221**, 495–507 (2012).

32. Merdith, A. S. et al. Extending full-plate tectonic models into deep time: linking the Neoproterozoic and the Phanerozoic. *Earth-Sci. Rev.* **214**, 103477 (2021).

33. Turner, S., Wilde, S., Wörner, G., Schaefer, B. & Lai, Y.-J. An andesitic source for Jack Hills zircon supports onset of plate tectonics in the Hadean. *Nat. Commun.* **11**, 1241 (2020).

34. Drabon, N. et al. Destabilization of long-lived Hadean protocrust and the onset of pervasive hydrous melting at 3.8 Ga. *AGU Adv.* **3**, e2021AV000520 (2022).

35. Solomatov, V. S. & Moresi, L. N. Three regimes of mantle convection with non-Newtonian viscosity and stagnant lid convection on the terrestrial planets. *Geophys. Res. Lett.* **24**, 1907–1910 (1997).

36. Moore, W. B. & Webb, A. A. G. Heat-pipe Earth. *Nature* **501**, 501–505 (2013).

37. Nutman, A. P. et al. Fifty years of the Eoarchean and the case for evolving uniformitarianism. *Precamb. Res.* **367**, 106442 (2021).

38. O'Neill, C., Marchi, S., Zhang, S. & Bottke, W. Impact-driven subduction on the Hadean Earth. *Nat. Geosci.* **10**, 793–797 (2017).

39. Spada, G., Ricard, Y. & Sabadini, R. Excitation of true polar wander by subduction. *Nature* **360**, 452–454 (1992).

40. Stern, R. A. in *Geological Survey of Canada, Current Research 1997-F. Radiogenic Age and Isotopic Studies: Report 10* 1–31 (Geological Survey of Canada, 1997).

41. Stern, R. A. & Amelin, Y. Assessment of errors in SIMS zircon U–Pb geochronology using a natural zircon standard and NIST SRM 610 glass. *Chem. Geol.* **197**, 111–146 (2003).

42. Davis, W. J., Pestaj, T., Rayner, N. & McNicoll, V. M. *Long-term reproducibility of 207Pb/206Pb age at the GSC SHRIMP lab based on the GSC Archean reference zircon z1242. Geological Survey of Canada, Scientific Presentation 111* (2019).

43. Bodorkos, S., Bowring, J. & Rayner, N. *Squid3: next-generation data processing software for Sensitive High Resolution Ion Micro Probe (SHRIMP). Exploring for the Future: Extended Abstracts, Geoscience Australia*(2020).

44. Steiger, R. H. & Jäger, E. Subcommission on geochronology: convention on the use of decay constants in geo- and cosmochronology. *Earth Planet. Sci. Lett.* **36**, 359–362 (1977).

45. Ludwig, K. R., *User's manual for Isoplot/Ex rev. 3.70: a geochronological toolkit for Microsoft Excel. Special Publication 4. Berkeley Geochronology Center* (2008).

46. Taylor, R. J. M. et al. Direct age constraints on the magnetism of Jack Hills zircon. *Sci. Adv.* **9**, eadd1511 (2023).

47. Lowe, D. R., Drabon, N., Byerly, G. R. & Byerly, B. L. Windblown Hadean zircons derived by erosion of impact-generated 3.3 Ga uplifts, Barberton Greenstone Belt, South Africa. *Precamb. Res.* **356**, 106–111 (2021).

48. Reddy, S. M., Timms, N. E., Pantleon, W. & Trimby, P. Quantitative characterization of plastic deformation of zircon and geological implications. *Contrib. Mineral. Petrol.* **153**, 625–645 (2007).

49. Fu, R. R. et al. Paleomagnetism of 3.5-4.0 Ga zircons from the Barberton Greenstone Belt, South Africa. *Earth Planet. Sci. Lett.* **567**, 116999 (2021).

50. Welch, B. L. The generalization of 'Student's' problem when several different population variances are involved. *Biometrika* **34**, 28–35 (1947).

51. Massey, F. J. The Kolmogorov-Smirnov test for goodness of fit. *J. Am. Stat. Assoc.* **46**, 68–78 (1951).

52. Mann, H. B. & Whitney, D. R. On a test of whether one of two random variables is stochastically larger than the other. *Ann. Math. Stat.* **18**, 50–60 (1947).

53. Merdith, A. S. et al. Extending full-plate tectonic models into deep time: linking the Neoproterozoic and the Phanerozoic. *Earth-Sci. Rev.* **214**, 103477 (2021).

54. Müller, R. D. et al. GPlates: building a virtual Earth through deep time. *Geochem. Geophys. Geosyst.* **19**, 2243–2261 (2018).

55. Aistleitner, C., Brauchart, J. S. & Dick, J. Point sets on the sphere \({{\mathbb{S}}}^{2}\) with small spherical cap discrepancy. *Discrete Comput. Geom.* **48**, 990–1024 (2012).

56. Serfling, R. J. Probability inequalities for the sum in sampling without replacement. *Ann. Stat.* **2**, 39–48 (1974).

57. Zegers, T. E., de Wit, M. J., Dann, J. & White, S. H. Vaalbara, Earth's oldest assembled continent? A combined structural, geochronological, and palaeomagnetic test. *Terra Nova* **10**, 250–259 (1998).

58. de Kock, M. O., Evans, D. A. D. & Beukes, N. J. Validating the existence of Vaalbara in the Neoarchean. *Precamb. Res.* **174**, 145–154 (2009).

59. Evans, M. E. & Muxworthy, A. R. Vaalbara palaeomagnetism. *Can. J. Earth Sci.* **56**, 912–916 (2019).

60. Cox, A. Lengths of geomagnetic polarity intervals. *J. Geophys. Res.* **73**, 3247–3260 (1968).

61. Merrill, R. T. & McFadden, P. L. Paleomagnetism and the nature of the geodynamo. *Science* **248**, 345350 (1990).

62. Tarduno, J. A., Cottrell, R. D. & Smirnov, A. V. High geomagnetic intensity during the mid-Cretaceous from Thellier analyses of single plagioclase crystals. *Science* **291**, 1779–1783 (2001).

63. Tarduno, J. A., Cottrell, R. D. & Smirnov, A. V. The Cretaceous superchron geodynamo: observations near the tangent cylinder. *Proc. Natl Acad. Sci. USA* **99**, 14020–14025 (2002).

64. Tarduno, J. A. & Cottrell, R. D. Dipole strength and variation of the time-averaged reversing and nonreversing geodynamo based on Thellier analyses of single plagioclase crystals. *J. Geophys. Res. Solid Earth* **110**, B11101 (2005).

65. Kulakov, E. V. et al. Analysis of an updated paleointensity database (Q_{PI}-PINT) for 65–200 Ma: implications for the long-term history of dipole moment through the Mesozoic. *J. Geophys. Res. Solid Earth* **124**, 9999–10022 (2019).

66. Glatzmaier, G. A., Coe, R. S., Hongre, L. & Roberts, P. H. The role of the Earth's mantle in controlling the frequency of geomagnetic reversals. *Nature* **401**, 885–890 (1999).

67. Hawkins, L. M. A. et al. Intensity of the Earth's magnetic field: evidence for a Mid-Paleozoic dipole low. *Proc. Natl Acad. Sci.* **118**, e2017342118 (2021).

Acknowledgements

We thank G. Kloc for the sample preparation, T. Pestaj in the SHRIMP laboratory and M. Polivchuk in the GSC field emission SEM laboratory. This work was performed in part at the University of Rochester Integrated Nanosystems Center (URnano), which is supported by the University of Rochester. A.H. thanks the Mpumalanga Tourism and Parks Agency and Sappi Forests for access and support. R.K.B. acknowledges support from the Leverhulme Early Career Fellowship (ECF-2020-617). This work was supported by the National Science Foundation (grant nos. EAR-1656348 and EAR-2051550 to J.A.T.). This is NRCan contribution number 20220385.

Author information

Authors and Affiliations

1. Department of Earth and Environmental Sciences, University of Rochester, Rochester, NY, USA

   John A. Tarduno, Rory D. Cottrell, Tinghong Zhou & Gautam Mitra

2. Department of Physics and Astronomy, University of Rochester, Rochester, NY, USA

   John A. Tarduno

3. Laboratory for Laser Energetics, University of Rochester, Rochester, NY, USA

John A. Tarduno

4. Geological Sciences, University of KwaZulu-Natal, Durban, South Africa

John A. Tarduno & Michael K. Watkeys

5. Geomagnetism Laboratory, University of Liverpool, Liverpool, UK

Richard K. Bono

6. Department of Earth, Ocean and Atmospheric Science, Florida State University, Tallahassee, FL, USA

Richard K. Bono

7. Natural Resources Canada, Geological Survey of Canada, Ottawa, Ontario, Canada

Nicole Rayner & William J. Davis

8. Department of Earth and Planetary Sciences, University of California, Santa Cruz, Santa Cruz, CA, USA

Francis Nimmo

9. Department of Geology, University of Johannesburg, Auckland Park, South Africa

Axel Hofmann & Jaganmoy Jodder

10. Evolutionary Studies Institute, University of the Witwatersrand, Wits, South Africa

Jaganmoy Jodder

11. Department of Geosciences, University of Arizona, Tucson, AZ, USA

Mauricio Ibañez-Mejia

12. Research Institute of Geology and Geoinformation, Geological Survey of Japan, National Institute of Advanced Industrial Science and Technology (AIST), Tsukuba, Japan

    Hirokuni Oda

Contributions

J.A.T. conceived and supervised the overall study. J.A.T., A.H. and J.J. conducted the field studies. M.I.-M. and R.D.C. separated zircons. R.D.C. conducted palaeomagnetic measurements and these data were analysed by R.D.C. and J.A.T. Reflected-light studies were conducted by R.D.C. and electron-microscope analyses by T.Z., J.A.T. and R.D.C. H.O. assisted in methods. Age data were collected by N.R. and analysed by N.R., W.J.D. and J.A.T. R.K.B. conducted statistical and plate-motion analyses. F.N. assisted in geodynamic interpretations, M.K.W. in tectonic implications and G.M. in microtectonic analysis. J.A.T. wrote the manuscript, with contributions from all the authors.

Corresponding author

Correspondence to John A. Tarduno.

Ethics declarations

Competing interests

The authors declare no competing interests.

Peer review

Peer review information

Nature thanks the anonymous reviewers for their contribution to the peer review of this work. Peer reviewer reports are available.

Additional information

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Extended data figures and tables

Extended Data Fig. 1 Silicate and apatite inclusions in BGS detrital zircons.

a, SEM secondary electron (SE2) images of zircon BGS2-z36 with apatite inclusion highlighted. **b**, Higher magnification of the inclusion with EDS analysis location highlighted. **c**, EDS spectra. **d**, Reflected-light microscope image of zircon BGS2-z39. **e**, SEM BSD image of zircon BGS2-z39 with quartz inclusion highlighted. **f**, Higher magnification of the inclusion with EDS analysis location highlighted. **g**, EDS spectra. **h**, Reflected-light microscope image of zircon BGS5-z30. **i**, SEM BSD image of zircon BGS5-z30 with feldspar inclusion identified. **j**, Higher magnification of the inclusion with EDS analysis locations highlighted. **k**–**n**, EDS spectra.

Source data

Extended Data Fig. 2 Fe-oxide and silicate inclusions in BGS detrital zircons.

a, SEM BSD image of zircon BGS2-z51 highlighting melt inclusions discussed (**b**,**f**,**d**); entire grain shown as inset with 50-µm scale bar. **b**, Higher-magnification SEM BSD image of feldspar inclusion b. EDS analysis location highlighted. **c**, EDS spectra of inclusion b. **d**, Higher-magnification SEM BSD image of multicomponent inclusion partially disrupted by polishing. EDS analysis location highlighted. **e**, EDS spectra from **d** suggesting a feldspar composition. **f**, Higher-magnification SEM BSD image of inclusion f (see Fig. 1 showing extinction with 90° change in polarization) showing EDS analysis locations. This inclusion may be partially disrupted by polishing. **g**, EDS spectra of main grain in **f** highlighting strong Fe signal. Al may be from adjacent feldspar (see **h**). **h**,

EDS of small grain in **f** showing spectra compatible with feldspar. **i**, Reflected-light microscopy image (100×) of zircon BGS2-z50. **j**, Reflected-light images (1,000× oil immersion) at 0° (top) and 90° (bottom) polarization showing extinction of Fe-oxide inclusion (magnetite) at depth (see Fig. 1). **k**, Reflected-light image of zircon BGS5-z36 (see Fig. 1). **l**, SEM BSD image of BGS5-z36. Fe particles shown in Fig. 1 are approximately 2 μm from the grain edge (also see 'Microtectonic analyses of zircons' section in Methods).

Source data

Extended Data Fig. 3 Further palaeointensity determinations and SHRIMP age data from individual BGS zircon crystals.

a, Concordia diagram showing SHRIMP geochronological analyses (uncertainty ellipses are 2σ) and palaeointensity value for zircon BGS1-z14. **b**, Corresponding backscatter scanning electron microscope with analysis spots labelled in **a**. **c**, Corresponding cathodoluminescence image. Images **b** and **c** shown with 50-μm scale. **d**–**f**, Analyses as shown in **a**–**c** for zircon BGS2-z17. **g**–**i**, Analyses as shown in **a**–**c** for zircon BGS5-z17. **j**–**l**, Analyses as shown in **a**–**c** for zircon BGS2-z10. Red, analyses from core; blue, analyses from rim.

Source data

Extended Data Fig. 4 BGS and JH zircon palaeointensity values versus time.

a, BGS palaeointensity data (green circles) and JH palaeointensity data (yellow circles), shown with Welch's *t*-test *P*-value (red curve; see 'Statistical analysis of BGS and JH zircon palaeointensity data' section in Methods). **b**, JH zircon palaeointensity values (Tarduno et al.[6,7]) with 100-Myr moving-window average. **c**, BGS zircon palaeointensity values shown with data of **a**. **d**, Residuals of JH palaeointensity values relative to model (JH 100-Myr moving-window average). **e**, Residuals of BGS palaeointensity values relative to model (JH 100-Myr moving-window average). **f**,

Empirical cumulative distribution function (ECDF) plot of JH and BGS residuals.

Source data

Extended Data Fig. 5 Plate-motion analysis 0–600 Ma.

a, Distribution of sites used in plate-motion analysis. Different colours distinguish between assigned plates. **b**, Representative example of the motion path for a single site with $\Delta\lambda$ at the median value (76°) determined using the plate-motion models described in the text. Solid line shows motion path resolved at 1-Myr intervals, black circles show 20-Myr steps. Blue circle, present-day site location; yellow circle, site palaeolocation at 600 Ma; red diamond, location at which the maximum latitudinal displacement is reached. **c**, Distribution of $\Delta\lambda$ from 228 sites located on 66 plates, determined using 1-Myr time steps shown with age of maximum latitudinal distance ($\Delta\lambda$) from present versus angle. **d**, Distribution of $\Delta\lambda$ in 5° bins; red line (and right *y* axis) shows empirical cumulative distribution of $\Delta\lambda$.

Source data

Extended Data Fig. 6 Maximum latitudinal motion ($\Delta\lambda$) of sites for the past 600 Myr.

Plate-motion-model-derived $\Delta\lambda$ values with (**a**) 1-Myr and (**b**) 100-Myr downsampling. Histograms show number of sites (total $n = 228$) with $\Delta\lambda$ binned into 5° groups. Solid vertical line shows median $\Delta\lambda$; dashed vertical lines mark the 95% interval for the distribution.

Source data

Extended Data Table 1 Palaeointensity determinations
Extended Data Table 2 BGS/JH comparisons: statistical parameters
Extended Data Table 3 100-Myr averages of JH, BGS and JH + BGS data

Supplementary information

Supplementary Information

This file contains Supplementary Text and Supplementary References

Peer Review File

Supplementary Table 1

SHRIMP U-Pb geochronology data

Source data

Source Data Fig. 1

Source Data Fig. 2

Source Data Fig. 3

Source Data Fig. 4

Source Data Extended Data Fig. 1

Source Data Extended Data Fig. 2

Source Data Extended Data Fig. 3

Source Data Extended Data Fig. 4

Source Data Extended Data Fig. 5

Source Data Extended Data Fig. 6

Rights and permissions

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by/4.0/.

Reprints and Permissions

About this article

Cite this article

Tarduno, J.A., Cottrell, R.D., Bono, R.K. *et al.* Hadaean to Palaeoarchaean stagnant-lid tectonics revealed by zircon magnetism. *Nature* **618**, 531–536 (2023). https://doi.org/10.1038/s41586-023-06024-5

- Received: 08 May 2022
- Accepted: 27 March 2023
- Published: 14 June 2023
- Issue Date: 15 June 2023
- DOI: https://doi.org/10.1038/s41586-023-06024-5

Share this article

Anyone you share the following link with will be able to read this content:

Get shareable link

Sorry, a shareable link is not currently available for this article.

Copy to clipboard

Provided by the Springer Nature SharedIt content-sharing initiative

---

This article was downloaded by **calibre** from https://www.nature.com/articles/s41586-023-06024-5

| Section menu | Main menu |

- Article
- Open Access
- Published: 07 June 2023

Unveiling the transition from niche to dispersal assembly in ecology

- Lynette H. L. Loke &
- Ryan A. Chisholm

Nature **volume 618**, pages 537–542 (2023)

- 7499 Accesses
- 103 Altmetric
- Metrics details

Abstract

A central goal in ecology is to understand what maintains species diversity in local communities. Classic ecological theory[1,2] posits that niches dictate the maximum number of species that can coexist in a community and that the richness of observed species will be below this maximum only where immigration is very low. A new alternative theory[3,4] is that niches, instead, dictate the minimum number of coexisting species and that the richness of observed species will usually be well above this because of ongoing immigration. We conducted an experimental test to discriminate between these two unified theories using a manipulative field experiment with tropical intertidal communities. We found, consistent with the new theory, that the relationship of species richness to immigration rate stabilized at a low value at low immigration rates and did not saturate at high immigration rates. Our results suggest that tropical intertidal communities have low niche

diversity and are typically in a dispersal-assembled regime where immigration is high enough to overfill the niches. Observational data from other studies[3,5] suggest that these conclusions may generalize to other ecological systems. Our new experimental approach can be adapted for other systems and be used as a 'niche detector' and a tool for assessing when communities are niche versus dispersal assembled.

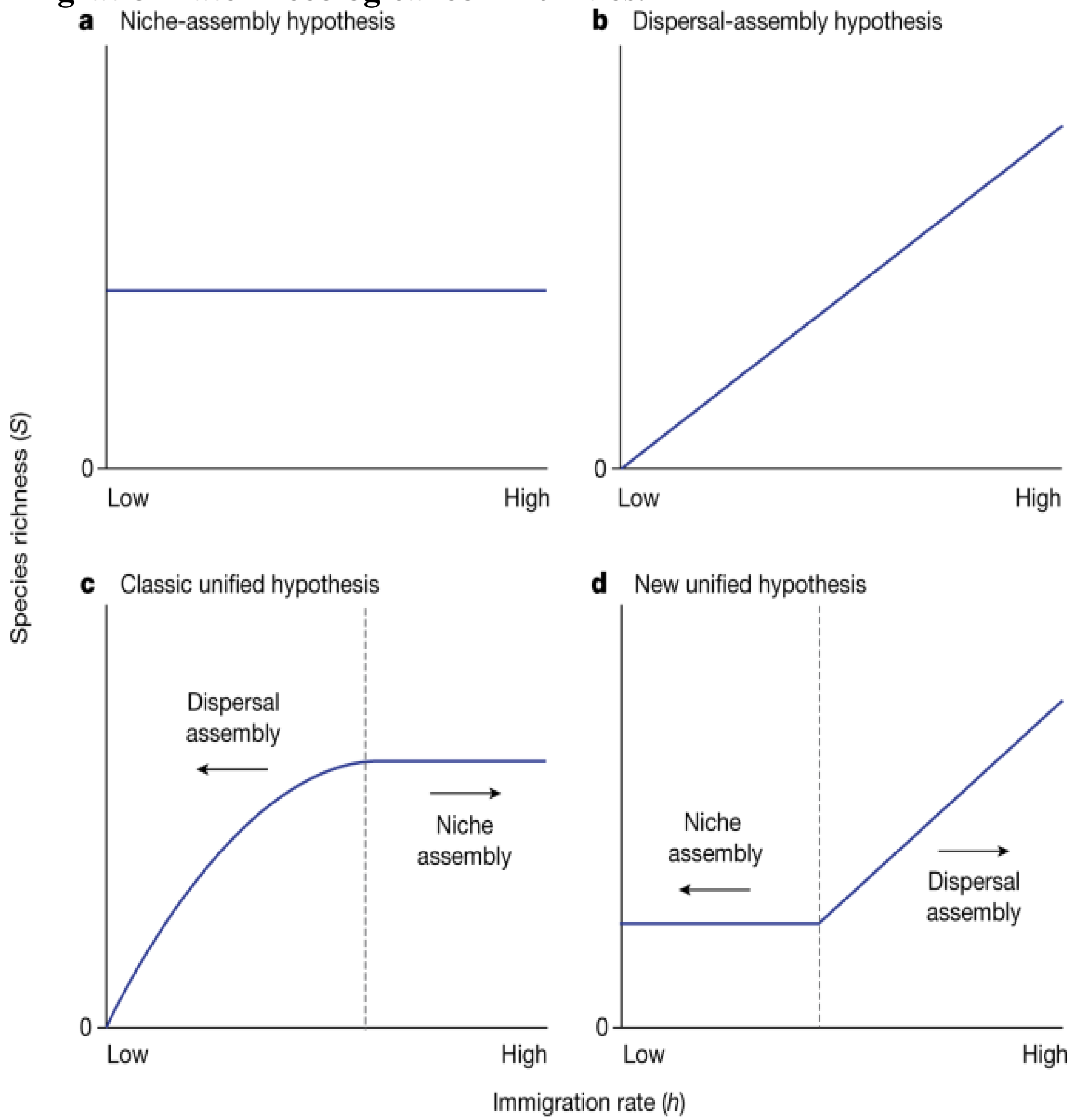
Main

A central unresolved question in ecology concerns the factors that drive and maintain species diversity in local communities. Under niche-assembly theory, species richness is explained in terms of stable coexistence attributable to properties of the local environment: species differ in their niches and these differences collectively stabilize community dynamics. Under dispersal-assembly theory, including MacArthur and Wilson's theory of island biogeography[1,2], the neutral theory of biodiversity[6] and mass effects[7], local diversity arises from an immigration–extinction balance and is ultimately dependent on regional processes.

Which of these two distinct perspectives is more correct is likely to depend on the context[8,9,10,11,12,13]. One critical mediating factor is the immigration rate to a community. Under pure niche assembly, there should be no relationship between species richness and the immigration rate (Fig. 1a). Under pure dispersal assembly, the relationship should be monotonically increasing (Fig. 1b). A classic unification of the two theories postulates that dispersal assembly applies only in cases where immigration is too low to fill all the niches (Fig. 1c). Indeed, MacArthur[14] and Wilson[15] speculated that the immigration–extinction balance and dispersal assembly central to their theory of island biogeography[1,2] applied only below some threshold immigration rate, above which species richness would be determined mainly by habitat diversity, that is, niche assembly[14] or by the diversity of the source species pool[15]. Under this hypothesis, the relationship between species richness and the immigration rate is predicted to increase for low immigration rates and then saturate at high immigration rates[15,16,17] (Fig. 1c). This saturation point reflects the maximum possible number of species

or the niche diversity of the island or patch—a ceiling that cannot be exceeded even with greater levels of immigration[14].

Fig. 1: Theoretical predictions for the relationship of species richness to immigration rate in ecological communities.



a, The niche-assembly hypothesis predicts constant species richness (\(S\)) with immigration rate (\(h\)). **b**, The dispersal-assembly hypothesis predicts a monotonically increasing relationship. **c**, A classic unifying hypothesis proposes that dispersal assembly applies under low immigration and niche assembly at high immigration. **d**, A new unifying hypothesis predicts the

opposite. The dashed lines in panels **c** and **d** represent the transition points between the two assembly regimes.

A recently formulated new unified hypothesis[3,4], however, predicts just the opposite: that niche assembly applies where immigration is low and dispersal assembly applies where immigration is high (Fig. 1d). Related ideas go back to Diamond's community assembly rules[18], the core–satellite hypothesis[19] and source–sink dynamics[20]. Under this hypothesis, niches stabilize diversity at some fixed level and provide a floor on species richness where immigration is low, but, as immigration increases, communities undergo a transition from a niche-assembled regime to a dispersal-assembled regime (Fig. 1d). The relationship of species richness to immigration rate is predicted to be nearly flat for low immigration rates and increasing for high immigration rates (Fig. 1d): that is, species richness asymptotes at low immigration rates rather than high immigration rates as in the classic hypothesis (Fig. 1c). The transition point occurs where immigration becomes sufficient to overwhelm the niche constraints, which leads to a dynamic immigration–extinction equilibrium with more species than niches.

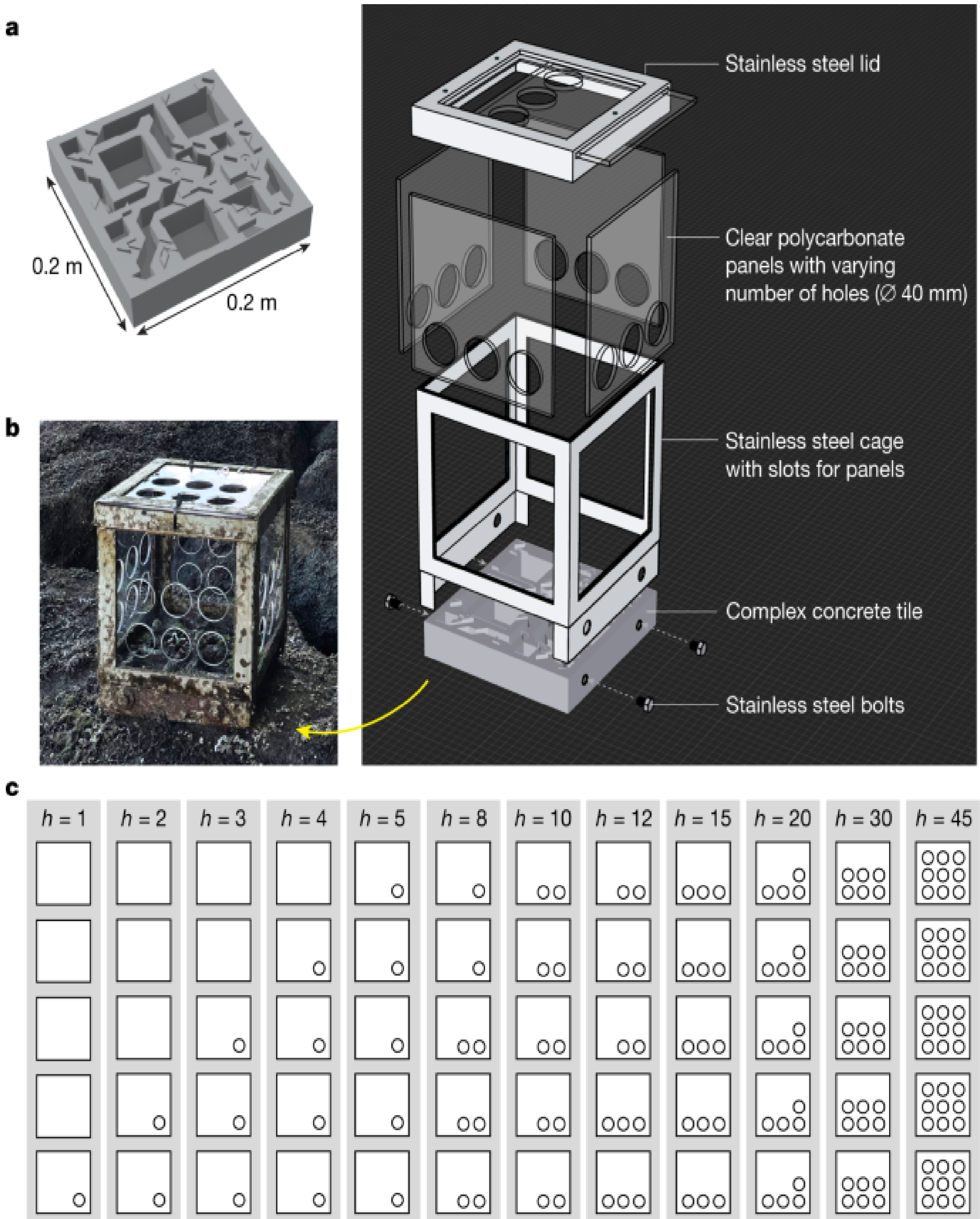
Which of these two contrasting unified hypotheses is correct has fundamental consequences for our understanding of the forces that structure ecological communities. The two hypotheses make very different predictions about the shape of the immigration–species richness curve, but there has not yet, to the best of our knowledge, been any experimental test to discriminate between them. A key challenge in experimentally testing the transition between dispersal- and niche-assembled regimes is the lack of foreknowledge as to the range over which the immigration rate should be varied to reveal both regimes (Fig. 1). In deterministic theoretical models, the niche-assembled regime can be uncovered by simply setting immigration to zero, but, in stochastic models or real experimental settings, this strategy fails because even in the niche-assembled regime, occasional immigrants are needed to offset stochastic extinctions. One feasible experimental approach is to vary the immigration rate across treatments with the rate approaching but not equal to zero in the low-immigration treatments. This is analogous to the mathematical approach of taking the limit as immigration tends to zero to avoid a singularity at zero. Ecologists acknowledge that immigration can in general influence community species richness[2,7,11], but the key to our

experimental approach is that we systematically vary immigration over a wide range in this way to reveal the functional form of the species richness versus immigration relationship and thus explain underlying mechanisms. Here, we use this approach to conduct a crucial experimental test of the predictions of the classic and new hypotheses to assess the conditions under which ecological communities are primarily niche versus dispersal assembled (Fig. 1).

Experimental design

We performed a manipulative field experiment on intertidal communities[21,22,23]. We focused on intertidal seawalls—a well-studied tractable model system[24,25,26] (Methods). We experimentally manipulated the level of immigration across experimental habitat patches using custom-built set-ups (Fig. 2). Each set-up was a treatment replicate with two components: (1) a topographically complex concrete habitat tile that provided a standardized level of niche diversity across all replicates (Fig. 2a and Methods); and (2) a stainless-steel cage, with clear square polycarbonate sheet panels (slotted into the cage), that fitted over each tile to create a seal to ensure that the immigration of organisms to and from each unit was only via fixed-size holes (40 mm diameter) cut into the clear panels (Fig. 2b and Methods). Each set-up was submerged at high tide (95.2% of the time) and exposed at low tide (4.8% of the time; see Methods). We created a total of 12 treatments by varying the number of holes (h) in the clear polycarbonate panels (Fig. 2c).

Fig. 2: Experimental set-up for manipulating the immigration rate to an intertidal community.



a,**b**, Each set-up comprised two components: a habitat tile, used to standardize habitat surface area and structural complexity across replicates (**a**); and a custom-built cage that fitted securely over the habitat tile to create a seal, with slots that allowed five clear polycarbonate sheet panels to be

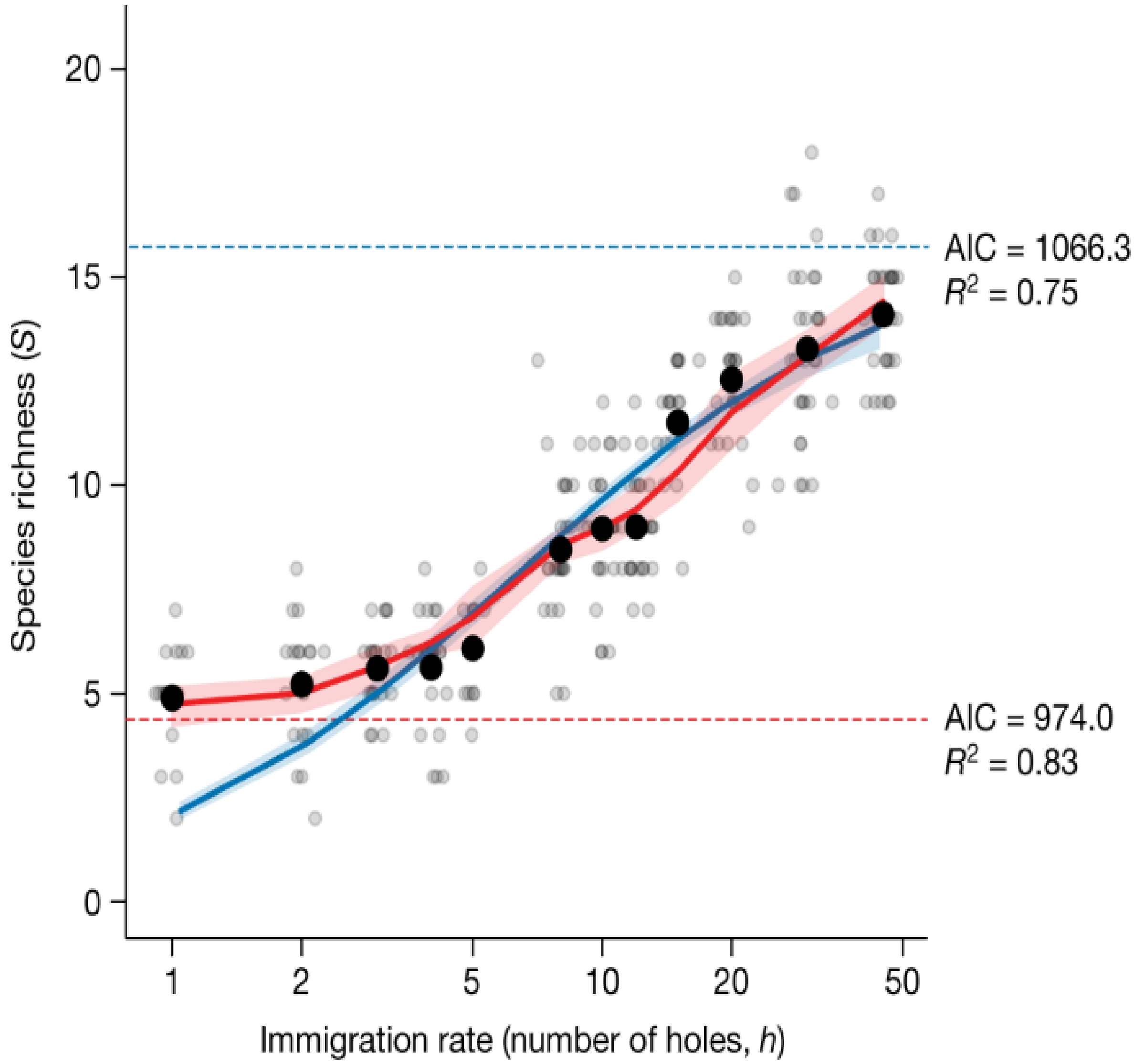
inserted and removed for regular maintenance (**b**). **c**, The level of immigration was controlled by varying the total number of holes *h* on the clear polycarbonate panels. We had 12 h-treatments (grey boxes; $h\,$= 1, 2, 3, 4, 5, 8, 10, 12, 15, 20, 30, 45) and five replicates ($n\,$ = 5) of each treatment set-up. The five white squares within each grey box show the number of holes and their arrangements for each of the five panels used to create the corresponding treatment. Panels in the first row were the top panels.

We installed a total of 60 experimental units (12 h-treatments × 5 replicates) in randomized order along the base of intertidal seawalls (approximately 0.5 m above chart datum) across a 400 m stretch at our study site in Singapore (Methods). To maintain the set-ups over the duration of the experiment, polycarbonate panels were removed and replaced with clean new ones every two to four weeks during low tide. We censused the number of species, species identities and abundances of the macrofaunal intertidal communities within each set-up every month for a year (Methods). Although we focused on macrofaunal intertidal species for practical reasons, we expect that qualitatively similar results would be obtained if we included algal and microbial communities. In total, we recorded 10,156 individuals of 64 different species: major taxa included gastropods, bivalves, polychaetes, tunicates and crustaceans (see full list in Supplementary Table 1). Species richness stabilized in all but the highest immigration treatments after six months (Extended Data Fig. 1), consistent with findings from previous studies carried out at the same study site and at nearby locations[25,26,27]. We thus used data from months 7 to 12 in our analyses (Extended Data Fig. 1). We fitted mechanistic mathematical models corresponding to each of the two hypotheses to the data by generalized nonlinear least squares (Methods). The classic model has two fitted parameters (the ratio of the immigration and extinction rates and metacommunity richness) and the new model has three (immigration rate parameter, niche diversity and the fundamental biodiversity number; see Methods). We assumed that immigration was linearly related to the number of holes (h), a parsimonious assumption that, if relaxed, only strengthens our results (Methods). We performed 1,000 bootstraps stratified by hole treatments to obtain 95% confidence intervals on the parameter estimates and fitted values (Methods).

Evidence of the transition

The relationship of species richness to immigration rate across experimental treatments was consistent with the new hypothesis, with a nearly flat phase at low immigration rates and an increasing phase at high immigration rates (Fig. 3; Akaike information criterion, AIC = 974.0). The fit of the classic model to the data was substantially poorer (Fig. 3, AIC = 1066.3; see also Extended Data Fig. 2, for a variant of the classic model). The new model explained 83% of the variation in the data. This provides experimental evidence of the transition from niche to dispersal assembly in ecology. The estimated niche diversity was ${n}^{* }=4.4$ with a 95% confidence interval of [4.2, 5.2], indicating that approximately four or five species can stably coexist in the seawall tile communities in the absence of substantial immigration, with species richness close to this minimum value being realized in the lowest immigration treatments ($h=1,\,2,\,3,\,4$ and 5). The transition to the dispersal-assembly regime occurred at intermediate immigration rates ($h\approx 10$) and roughly three times the minimum number of species were present in the highest immigration treatment ($h=45$). The effect of increasing species richness was partly mediated by increasing total abundance with immigration among the high-immigration treatments and partly by increasing species richness given total abundance (Fig. 4).

Fig. 3: Evidence of the transition from niche to dispersal assembly.

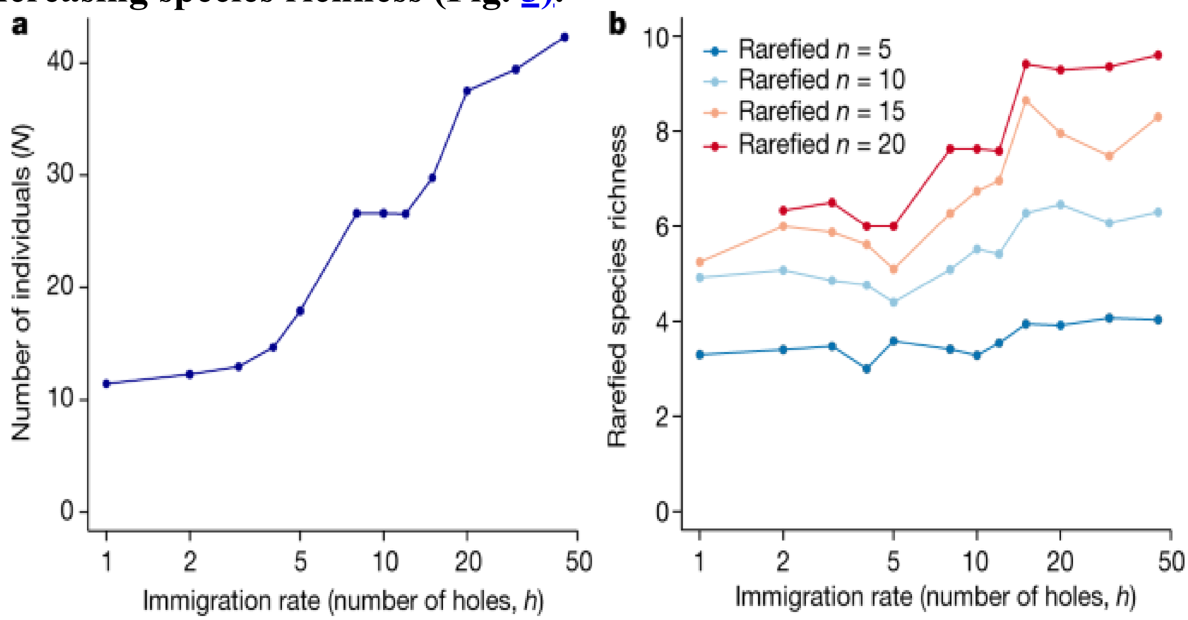


Each small grey point indicates a single experimental unit (seawall tile; 12 \(h\)-treatments, \(n\,\) = 5) in a given month (*x* coordinates are jittered to improve visibility); each large black point shows a treatment mean. Fitted models are shown by the solid curves (mean estimates) with shaded 95% confidence intervals from 1,000 bootstraps; the new model (equation (6)) is shown in red whereas the classic model is in blue (equation (3); see Methods). Consistent with the new model, the relationship of species richness to immigration in experimental intertidal seawall communities is biphasic and asymptotes to a low value at low immigration rates, indicating that, for low immigration rates, communities are primarily niche assembled and comprise approximately four or five species (dashed red horizontal line) and that, for high immigration rates, communities are primarily dispersal

assembled. Under the new model, hypothetical species richness due to niche assembly alone is constant and close to richness in the low-immigration treatments (dashed red horizontal line). The classic model was substantially poorer than the new model (judged by Akaike information criterion, AIC). The dashed blue horizontal line indicates the expected species richness at high levels of immigration under the classic model, as estimated by the fitted model.

Source Data

Fig. 4: As immigration is increased, both increasing number of individuals and increasing rarefied species richness contribute to increasing species richness (Fig. 3).



a, Average number of individuals (N) for each value of immigration (h). **b**, Average rarefied species richness for each value of immigration (h), using various rarefaction sample sizes (n).

Source Data

Discussion

Our experimental results are consistent with the new hypothesis, that is, that communities are niche assembled under low immigration and are otherwise

primarily dispersal assembled[3,4]. This is consistent with previous observational evidence of a biphasic species–area relationship in island archipelagos, where area is interpreted as a proxy for immigration[3]. However, the island data, being observational, are subject to alternative interpretations[5] because confounding variables, such as niche diversity, can covary with area[7]. Our experimental approach avoids this by standardizing area across all treatments and thus provides a more direct test of theoretical principles about the effects of immigration on community assembly. We note that species richness was still increasing slowly and did not fully equilibrate in our highest immigration treatments ($h=30,\,45$) after six months (Extended Data Fig. 1): allowing it to equilibrate would only strengthen our conclusions by accentuating the biphasic relationship of species richness to immigration rate (species richness would be higher towards the right of Fig. 3).

Our new experimental approach can be adapted to other systems such as annual plant communities[28] and microbial communities[29] in laboratory conditions. In addition to providing additional tests of our two central hypotheses (Fig. 1c,d), the experimental approach can be used to assess when communities are niche versus dispersal assembled—thereby shedding light on a question that has perplexed ecologists for decades. The experiments also act as a 'niche detector', that is, a tool for assessing how many species can coexist without substantial immigration—another long-standing ecological conundrum.

Our experimental data, taken together with past observational and theoretical studies, suggest that niche diversity in ecological systems is typically low. The estimated niche diversity of our intertidal seawall model system was low (${n}^{*}=4.4$) and we predict that similar values would be obtained in experiments on the natural analogues of our system, tropical rocky shores. Low niche diversity for a variety of ecological communities is also implied by island species–area relationships, which typically exhibit a similar shape to our species–immigration relationship (Fig. 3), with a flat low-richness phase at low island areas attributable to niche assembly under low immigration[3]. This flat phase has long been known as the 'small-island effect'[30,31], but lacked a satisfactory dynamical explanation until recently[3]. Theoretical models also suggest that the number of species that can coexist

via local niche mechanisms is generally low. Although, in some theoretical models, a large number of species can stably coexist via temporal niches arising from relative nonlinearity[32,33] and storage effects[34], such coexistence is fragile and not robust to stochasticity and perturbations, which raises doubt about its relevance in nature[35].

We predict that applications of our experimental approach to other systems will lend further support to the hypothesis that niche diversity is generally much lower than the typical total species richness of natural habitats, which implies that most species are not stably coexisting but, instead, are transiently co-occurring and are reliant on continued immigration[8] (that is, mass effects[7]). A corollary is that most natural communities are in the dispersal-assembly regime. In rocky intertidal communities characteristic of our study site, species richness is invariably much higher than observed in our niche-assembly regime[36,37]. We conjecture that these conclusions also apply to highly diverse systems such as tropical rainforest tree communities, where hundreds of species may be present in a single hectare[38]. A prominent niche-assembly explanation for high tree species richness is that predators and pathogens are largely host-specific, which leads to a high diversity of enemy-escape niches[39]. By contrast, the dispersal-assembly explanation attributes high local tree species richness to immigration, as well as to potentially larger-scale niche processes (for example, source–sink dynamics[20], with species being adapted to distinct habitats across the landscape) and, ultimately, to diversification processes that play out over geological timescales and continental spatial scales. Although our experimental approach would be impractical in tropical forests because of long tree generation times, the question of what determines high tree diversity could be indirectly informed by evidence from other, more experimentally tractable, high-diversity communities, such as plankton communities[33,40]. If plankton communities collapse to a small number of species under low immigration, similar to our seawall communities (Fig. 3), this would cast doubt on the general notion that an observation of high species diversity implies high niche diversity.

Given the evidence for two distinct assembly regimes, we encourage future experimental work—whether on intertidal systems or other communities—to carefully distinguish which regime prevails in any given context and to

study each regime separately. In the niche-assembly regime, although we observed some consistency in the set of coexisting species (for example, the gastropod *Drupella margariticola* and the tunicate *Didemnum psammatodes*), our experimental set-up was not designed to unravel the precise niche-assembly rules. Figuring out the exact niche mechanisms operating in this regime is a priority for future research. Experiments that directly manipulate the species identities of immigrants can potentially address this (as done in microbial communities[29]). An alternative would be to spread the experimental units over a much larger geographical area to capture a larger species pool. For our study system, our estimate of the transition point between the two regimes ($h\approx 10$) can inform the design of future studies focused on one regime or the other. Our results can also inform sample size determination for future studies focusing on more-detailed biodiversity patterns including co-occurrence matrices and species abundance distributions. One assumption of our method is consistency of abiotic conditions across experimental immigration treatments. We confirmed this for light and temperature (Methods), but it remains conceivable that other variables (for example, pH) are important and we encourage future similar experiments to measure as wide a range of abiotic variables as possible. More broadly, to understand the dispersal-assembly regime, ecologists must design more cross-scale studies that link local and regional dynamics: it is increasingly clear that the factors driving the diversity of a small local community cannot be unravelled by studying the system in isolation, that is, without accounting for immigration[9,11]. A holistic assessment of the dispersal-assembly regime should also account for niche processes that may not on their own permit stable coexistence (as in the niche-assembly regime) but nevertheless slow competitive exclusion and thus act synergistically with immigration to increase species richness[41,42]. Such synergies between immigration and niche processes may turn out to be more important for diversity than traditional stable coexistence.

Our results have practical implications for the use of concrete tiles to augment diversity on man-made coastal defences in many parts of the world—a strategy known as eco-engineering or nature-based solutions[43,44]. The predominant role of dispersal assembly in our system indicates that such efforts may be ineffective unless attention is also paid to habitat preservation and connectivity in the broader landscape to ensure a large, diverse,

continued supply of immigrants. Regardless of their topographic complexity, concrete tiles placed in a depauperate intertidal landscape will harbour few species and this needs to be considered more explicitly, especially in real-world applications. More broadly, the importance of dispersal assembly and the apparent low number of stabilized niches cautions against conservation approaches that rely too strongly on niche mechanisms for preserving species diversity[45]. Efforts to protect biodiversity must account for large-scale landscape connectivity to be successful—although we encourage further experimental tests of the new hypothesis in other systems and locations to establish the generality of this conclusion.

The emerging picture, not only from our experimental study but also from observations on island archipelagos and from theoretical work, is that dispersal assembly is the prevailing regime for most ecological communities and that the number of stabilized niches is low. This may explain the relative success of niche theory in low-diversity experimental systems[28,46,47] but its failure to provide a predictive framework in most natural systems, especially in high-diversity systems such as tropical rainforests and coral reefs[48,49,50]. The quest for a general understanding of biological diversity would benefit from more experimental tests of which assembly regime—niche or dispersal—prevails in different contexts[48].

Methods

Study system

We conducted the study on granite rip-rap seawalls in Singapore. Singapore is a small (728.6 km²) tropical island city-state in Southeast Asia: it comprises a mainland (approximately 710 km²), which is approximately 50 km from east to west and 26 km from north to south, and 64 smaller offshore islands, including Sentosa island where the present study was conducted (specifically, at 1° 15′ 00.4″ N, 103° 49′ 04.0″ E). More than 80% of Singapore's natural coastline has been modified and replaced by granite rip-rap seawalls[51]. Natural rocky shores are now limited to a short 300-m stretch along the southern shore and a few islands south of Singapore's mainland (including parts of Sentosa island). Located just one degree above the

equator, our study system experiences a typical tropical climate: high and uniform temperatures, abundant rainfall and high humidity all year round. For the duration of the study, from March 2021 to March 2022, the average reported temperature was 27.9 °C and the average monthly precipitation was 19.6 cm at our study site[52]. Facing the Singapore Strait, our study system is an open hard-bottomed system in a highly urbanized marine environment that experiences a mixed semi-diurnal tidal regime[51,53].

Monthly biodiversity surveys over a year conducted in ref. [36], which included our study location, found that, whereas composition of the communities colonizing seawalls and rocky shores differed, the two habitats harboured similar suites of species. This pattern is consistent with hard-bottom benthic systems around the world, including those in temperate and sub-tropical regions[24,54,55]. The study[36] also found no significant temporal patterns in species assemblages (that is, our study system can be considered aseasonal), with molluscs, crustaceans and algae dominating all year round (see also ref. [56]). Although rocky shores often host greater species richness compared with seawalls, many of the species unique to either habitat are rare species: for instance, in their monthly surveys over a year, ref. [36] found less than ten individuals for all unique species in either habitat. What is notable is that, although seawalls are artificial, in the tropics they can still harbour relatively high species numbers: in Singapore, the authors of ref. [36] found a total of 138 species on natural rocky shores versus 105 species on seawalls through their year-long survey. Communities on aseasonal tropical hard shores also undergo relatively rapid rates of succession[27,57,58] compared with temperate and seasonal tropical hard shores[59,60].

The seawalls at our study location were constructed from boulders approximately 50 years ago and are not grouted with concrete or any other fill material[61]. These boulders are not regularly replaced or maintained, due to the ideal strength and material properties afforded by granite and the relatively low wave energy in the meso-tidal coastal environment of Singapore[51,62,63]. The biota on these hard structures, which act as the source pool for immigrants to our experimental units, can thus be considered climax communities[24,57].

Habitat tile design and fabrication

Concrete tiles (Fig. 2a) measuring 0.2 m × 0.2 m × 0.06 m (width × length × depth) were used as habitat units from which we sampled our ecological communities. This approach is advantageous as it allows us to standardize habitat area and niche diversity across all replicates. Only one tile design was used and it comprised two microhabitat component types relevant to intertidal organisms: square pits and grooves[25,26,64]. These structural components varied in their sizes and depths from 0.01 m to 0.04 m; specifically, pits were first positioned randomly on the tile, after which we overlaid the remaining unoccupied space with grooves using a Truchet tiling algorithm. This was done using the three-dimensional modelling software Rhino 6. Once the tile design was finalized, the model was three-dimensional printed and from this rubber moulds were made. The actual experimental tiles were then cast from the moulds using concrete (1:3 cement to sand mix)[25]. During casting, four stainless-steel M8 hex nuts were embedded into the tile to act as insets for attaching the stainless-steel cage set-up to each tile unit using stainless-steel bolts (Fig. 2b). Temperature and light conditions in the experimental units were comparable with the surrounding seawall (Supplementary Fig. 1). Previous studies conducted worldwide[65] and in Singapore (including our study location; for example, refs. [25,26,27,57,58]) have shown tiles to be an effective experimental device for sampling the benthic macrofaunal diversity of these tropical intertidal communities.

Determining diameter of hole on panels

To determine the diameter of holes on panels used in the experimental set-up (Fig. 2), we conducted a field survey of the study site (before the start of the experiment, in February 2021), as well as an adjacent natural rocky shore site located further along the same shoreline approximately 1 km away. Five 1 m^2 quadrats were placed randomly along a 100 m transect at the mid-tidal height (approximately 0.5 m above chart datum; that is, along the same tidal height as in our experiment) at each site. From each quadrat, we counted the number of individuals of all macrofauna (greater than 1 mm) found within the 1 m^2 quadrat and measured the maximum body length of each individual using a digital Vernier caliper (with a resolution of 0.01 mm). A total of 245

individuals were recorded. The mean maximum body length of individuals was 16.8 mm (the mean was 15.0 mm for seawalls and 18.4 mm for rocky shores; Supplementary Fig. 2). Only two of the 245 individuals (less than 1%) had maximum body lengths greater than 40 mm (these two were gastropod species: *Chicoreus capucinus* (a whelk), measuring 42.8 mm, and *Peronia verruculata* (an onch slug), measuring 40.4 mm). Note, however, that their widths were less than 40 mm and that both of these species were present in the actual experimental communities.

Data analyses

We fitted formulas arising from mechanistic mathematical models corresponding to each of the two hypotheses to the data on species richness *S* versus number of holes *h*, using generalized nonlinear least squares (function gnls in the R programming language[66]). The classic model formula arises from a dynamic model describing how the rate of change of species richness on an island depends on the immigration and extinction rates, yielding the equilibrium solution[15,16,67].

$$S=\frac{\lambda P}{\lambda +\mu }$$
(1)

where P is the number of species in the metacommunity, λ is the per-species immigration rate and μ is the per-species extinction rate. In the first version of this model[15], the extinction rate is assumed to be inversely proportional to the area of the island ($\mu ={\mu }^{{\prime} }/A$), giving

$$S=\frac{\lambda P}{\lambda +\frac{{\mu }^{{\prime} }}{A}}$$
(2)

In our study system, area A is constant and we assume that the immigration rate is proportional to the number of holes h, so that $\lambda =\alpha h$ for some parameter α. Writing ${\alpha }^{{\prime} }=\alpha A/{\mu }^{{\prime} }$ then gives a two-parameter relationship between S and h:

$$S=\frac{{\alpha }^{{\prime} }hP}{{\alpha }^{{\prime} }h+1}$$

(3)

In the second version of this model[16,67], the extinction rate is, instead, assumed to be proportional to the number of species currently on the island ($\mu ={\mu }^{{\prime\prime} }S$), giving

$$S=\frac{\lambda P}{\lambda +{\mu }^{{\prime\prime} }S}$$
$$\Rightarrow S=\frac{-\lambda +\sqrt{\lambda \left(\lambda +4P{\mu }^{{\prime\prime} }\right)}}{2{\mu }^{{\prime\prime} }}$$
(4)

Again setting $\lambda =\alpha h$ and writing ${\alpha }^{{\prime\prime} }=\alpha /{\mu }^{{\prime\prime} }$, gives another two-parameter relationship between S and h:

$$S=\frac{-{\alpha }^{{\prime\prime} }h+\sqrt{{\alpha }^{{\prime\prime} }h\left({\alpha }^{{\prime\prime} }h+4P\right)}}{2}$$
(5)

Under the alternative new hypothesis, a mathematical formula for the relationship of species richness to number of holes arises from a master equation describing neutral drift within each of a fixed number of equal-sized non-overlapping niches in a local community[68,69]. Occasional stochastic extinction is balanced by immigration of new individuals from a larger metacommunity. The model has four parameters: the number of niches ${n}^{* }$, the local community size J, the probability that a new individual is an immigrant m and the fundamental biodiversity number θ of the metacommunity from which immigrants are drawn. Species richness at the dynamic equilibrium is[3]

$$S\approx \theta \left\{{\psi }_{0}\left(\frac{\theta }{{n}^{* }}+\,{\gamma }^{* }\left({\psi }_{0}\left({\gamma }^{* }+{J}^{* }\right)-\,{\psi }_{0}\left({\gamma }^{* }\right)\right)\right)-{\psi }_{0}\left(\frac{\theta }{{n}^{* }}\right)\,\right\}$$
(6)

where ${\psi }_{0}$ is the digamma function[70], ${\gamma }^{* }=(\,{J}^{* }-1)m/(1-m)$ is a composite immigration parameter and \({J}^{*

}=J/{n}^{* }\) is the number of individuals in each niche. The original derivation of equation (6) assumed death–birth dynamics, where a death event leads to a vacated space that can be filled by the birth of a new individual, but the same equilibrium solution arises from birth–death dynamics where a newly born individual displaces and kills an established individual. Other variants of the standard model, for example, with abundance measured as biomass instead of number of individuals, lead to similar equilibrium solutions. To relate the composite immigration parameter ${\gamma }^{* }$ to the number of holes in our seawall experiment, we assume that the number of immigrants per unit time is proportional to the number of holes and that the number of local births in a given niche is proportional to the number of individuals residing there (minus one for the individual that just died to vacate a space), so that the probability of immigration is $m=bh/(bh+c(\,{J}^{* }-1))$ for some constants b and c, which leads to ${\gamma }^{* }=bh/(c{n}^{* })$. We then define the composite parameter $\alpha =b/(c{n}^{* })$ to get ${\gamma }^{* }=\alpha h$. Given data on J, the model then has three free parameters, $\alpha ,\,{n}^{* }$ and θ, which can be estimated from data on S and h.

Note that in the new model the per-individual immigration rate scales linearly with number of holes, but in the classic model it is the per-species immigration rate that scales in this way. We argue that the novel model is more realistic because the per-species rate is likely to exhibit a convex relationship with number of holes (similar to a species accumulation curve). Modifying the classic model to allow a convex relationship would be likely to reduce the quality of its fit to the data (because it would be harder for it to capture the observed upswing in species richness at high numbers of holes), but we chose not to do this because the linear relationship is consistent with Wilson's original conceptualization[15] and because we wanted to conduct a conservative test of the new hypothesis.

We fitted each of the three models (two versions of the classic model, equations (3) and (5); and the new model, equation (6)) to the data by generalized least squares. After fitting each model, we performed 1,000 bootstraps stratified by hole treatments to obtain 95% confidence intervals on the parameter estimates and fitted values of species richness. The

distribution of errors for each model was close to Gaussian in all cases (Supplementary Fig. 3).

Reporting summary

Further information on research design is available in the Nature Portfolio Reporting Summary linked to this article.

Data availability

The data generated in this study are available in the Zenodo repository: https://doi.org/10.5281/zenodo.7819940. Source data are provided with this paper.

Code availability

The R code used for data analysis and for producing the figures are available in the Zenodo repository: https://doi.org/10.5281/zenodo.7819940.

References

1. MacArthur, R. H. & Wilson, E. O. An equilibrium theory of insular zoogeography. *Evolution* **17**, 373–387 (1963).

2. Macarthur, R. H. & Wilson, E. O. *The Theory of Island Biogeography* (Princeton Univ. Press, 1967).

3. Chisholm, R. A., Fung, T., Chimalakonda, D. & O'Dwyer, J. P. Maintenance of biodiversity on islands. *Proc. R. Soc. B: Biol. Sci.* **283**, 20160102 (2016).

4. Chisholm, R. A. & Fung, T. Examining the generality of the biphasic transition from niche-structured to immigration-structured communities. *Theor. Ecol.* **15**, 1–16 (2022).

5. Schrader, J., Moeljono, S., Keppel, G. & Kreft, H. Plants on small islands revisited: the effects of spatial scale and habitat quality on the species–area relationship. *Ecography* **42**, 1405–1414 (2019).

6. Hubbell, S. P. *The Unified Neutral Theory of Biodiversity and Biogeography* (Princeton Univ. Press, 2001).

7. Shmida, A. V. I. & Wilson, M. V. Biological determinants of species diversity. *J. Biogeogr.* **12**, 1–20 (1985).

8. Leibold, M. A. & McPeek, M. A. Coexistence of the niche and neutral perspectives in community ecology. *Ecology* **87**, 1399–1410 (2006).

9. Chase, J. M. & Myers, J. A. Disentangling the importance of ecological niches from stochastic processes across scales. *Philos. Trans. R. Soc. B* **366**, 2351–2363 (2011).

10. Chase, J. M. et al. Embracing scale-dependence to achieve a deeper understanding of biodiversity and its change across communities. *Ecol. Lett.* **21**, 1737–1751 (2018).

11. Leibold, M. A. et al. The metacommunity concept: a framework for multi-scale community ecology. *Ecol. Lett.* **7**, 601–613 (2004).

12. Tilman, D. Niche tradeoffs, neutrality, and community structure: a stochastic theory of resource competition, invasion, and community assembly. *Proc. Natl Acad. Sci. USA* **101**, 10854–10861 (2004).

13. Kadmon, R. & Allouche, O. Integrating the effects of area, isolation, and habitat heterogeneity on species diversity: a unification of island biogeography and niche theory. *Am. Nat.* **170**, 443–454 (2007).

14. MacArthur, R. H. Patterns of species diversity. *Biol. Rev.* **40**, 510–533 (1965).

15. Wilson, E. O. The species equilibrium. *Brookhaven Sym. Biol.* **22**, 38–47 (1969).

16. Wright, S. J. Intra-archipelago vertebrate distributions: the slope of the species-area relation. *Am. Nat.* **118**, 726–748 (1981).

17. Lomolino, M. V. & Weiser, M. D. Towards a more general species-area relationship: diversity on all islands, great and small. *J. Biogeogr.* **28**, 431–445 (2001).

18. Diamond, J. M. in *Ecology and Evolution of Communities* (eds Cody, M. L. & Diamond, J. M.) 342–444 (Harvard Univ. Press, 1975).

19. Hanski, I. Dynamics of regional distribution: the core and satellite species hypothesis. *Oikos* **38**, 210–221 (1982).

20. Pulliam, H. R. Sources, sinks, and population regulation. *Am. Nat.* **132**, 652–661 (1988).

21. Paine, R. T. & Vadas, R. L. The effects of grazing by sea urchins, *Strongylocentrotus* spp., on benthic algal populations 1. *Limnol. Oceanogr.* **14**, 710–719 (1969).

22. Lubchenco, J. & Menge, B. A. Community development and persistence in a low rocky intertidal zone. *Ecol. Monogr.* **48**, 67–94 (1978).

23. Bertness, M. D., Leonard, G. H., Levine, J. M., Schmidt, P. R. & Ingraham, A. O. Testing the relative contribution of positive and negative interactions in rocky intertidal communities. *Ecology* **80**, 2711–2726 (1999).

24. Hawkins, S. J., Pack, K. E., Hyder, K., Benedetti-Cecchi, L. & Jenkins, S. R. Rocky shores as tractable test systems for experimental ecology. *J. Mar. Biol. Assoc. UK* **100**, 1017–1041 (2020).

25. Loke, L. H. L. & Todd, P. A. Structural complexity and component type increase intertidal biodiversity independently of area. *Ecology* **97**, 383–393 (2016).

26. Loke, L. H. L., Chisholm, R. A. & Todd, P. A. Effects of habitat area and spatial configuration on biodiversity in an experimental intertidal

community. *Ecology* **100**, e02757 (2019).

27. Hartanto, R. S. et al. Material type weakly affects algal colonisation but not macrofaunal community in an artificial intertidal habitat. *Ecol. Eng.* **176**, 106514 (2022).

28. Levine, J. M. & HilleRisLambers, J. The importance of niches for the maintenance of species diversity. *Nature* **461**, 254–257 (2009).

29. Friedman, J., Higgins, L. M. & Gore, J. Community structure follows simple assembly rules in microbial microcosms. *Nat. Ecol. Evol.* **1**, 1–7 (2017).

30. Triantis, K. A. & Sfenthourakis, S. Island biogeography is not a single-variable discipline: the small island effect debate. *Divers. Distrib.* **18**, 92–96 (2012).

31. Preston, F. W. The canonical distribution of commonness and rarity: part I. *Ecology* **43**, 185–215 (1962).

32. Armstrong, R. A. & McGehee, R. Competitive exclusion. *Am. Nat.* **115**, 151–170 (1980).

33. Huisman, J. & Weissing, F. J. Biodiversity of plankton by species oscillations and chaos. *Nature* **402**, 407–410 (1999).

34. Chesson, P. Mechanisms of maintenance of species diversity. *Annu. Rev. Ecol. Syst.* **31**, 343–366 (2000).

35. Schippers, P., Verschoor, A. M., Vos, M. & Mooij, W. M. Does "supersaturated coexistence" resolve the "paradox of the plankton"? *Ecol. Lett.* **4**, 404–407 (2001).

36. Lai, S., Loke, L. H. L., Bouma, T. J. & Todd, P. A. Biodiversity surveys and stable isotope analyses reveal key differences in intertidal assemblages between tropical seawalls and rocky shores. *Mar. Ecol. Prog. Ser.* **587**, 41–53 (2018).

37. Lim, L. J. W. et al. Diversity and distribution of intertidal marine species in Singapore. Singapore. *Raffles Bull. Zool.* **68**, 396–403 (2020).

38. Turner, I. M. *The Ecology of Trees in the Tropical Rain Forest* (Cambridge Univ. Press, 2001).

39. Terborgh, J. Using Janzen–Connell to predict the consequences of defaunation and other disturbances of tropical forests. *Biol. Conserv.* **163**, 7–12 (2013).

40. Descamps-Julien, B. & Gonzalez, A. Stable coexistence in a fluctuating environment: an experimental demonstration. *Ecology* **86**, 2815–2824 (2005).

41. Levi, M. R. & Bestelmeyer, B. T. Digital soil mapping for fire prediction and management in rangelands. *Fire Ecol.* **14**, 1–12 (2018).

42. Chisholm, R. A. & Fung, T. Janzen-Connell effects are a weak impediment to competitive exclusion. *Am. Nat.* **196**, 649–661 (2020).

43. Morris, R. L. et al. Design options, implementation issues and evaluating success of ecologically engineered shorelines. *Oceanogr. Mar. Biol.* **57**, 169–228 (2019).

44. Cohen-Shacham, E., Walters, G., Janzen, C. & Maginnis, S. *Nature-based Solutions to Address Global Societal Challenges* (IUCN, Gland, Switzerland, 2016).

45. Cordonnier, T., Kunstler, G., Courbaud, B. & Morin, X. Managing tree species diversity and ecosystem functions through coexistence mechanisms. *Ann. For. Sci.* **75**, 1–11 (2018).

46. Tilman, D. Tests of resource competition theory using four species of Lake Michigan algae. *Ecology* **62**, 802–815 (1981).

47. Fargione, J., Brown, C. S. & Tilman, D. Community assembly and invasion: an experimental test of neutral versus niche processes. *Proc. Natl Acad. Sci. USA* **100**, 8916–8920 (2003).

48. Hubbell, S. P. Neutral theory in community ecology and the hypothesis of functional equivalence. *Funct. Ecol.* **19**, 166–172 (2005).

49. Volkov, I., Banavar, J. R., Hubbell, S. P. & Maritan, A. Patterns of relative species abundance in rainforests and coral reefs. *Nature* **450**, 45–49 (2007).

50. Dornelas, M., Connolly, S. R. & Hughes, T. P. Coral reef diversity refutes the neutral theory of biodiversity. *Nature* **440**, 80–82 (2006).

51. Lai, S., Loke, L. H. L., Hilton, M. J., Bouma, T. J. & Todd, P. A. The effects of urbanisation on coastal habitats and the potential for ecological engineering: a Singapore case study. *Ocean Coast. Manage.* **103**, 78–85 (2015).

52. Climate of Singapore. *Meteorological Service Singapore* http://www.weather.gov.sg/climate-climate-of-singapore/ (2022).

53. Van Maren, D. S. & Gerritsen, H. Residual flow and tidal asymmetry in the Singapore Strait, with implications for resuspension and residual transport of sediment. *J. Geophys. Res.* **117**, C04021 (2012).

54. Chapman, M. G. & Bulleri, F. Intertidal seawalls—new features of landscape in intertidal environments. *Landsc. Urban Plan.* **62**, 159–172 (2003).

55. Davis, J., Levin, L. & Walther, S. Artificial armored shorelines: sites for open-coast species in a southern California bay. *Mar. Biol.* **140**, 1249–1262 (2002).

56. Lee, A. C. & Sin, T. M. Intertidal assemblages on coastal defence structures in Singapore II. Contrasts between islands and the mainland. *Raffles Bull. Zool.* **22**, 255–268 (2009).

57. Loke, L. H. L., Liao, L. M., Bouma, T. J. & Todd, P. A. Succession of seawall algal communities on artificial substrates. *Raffles Bull. Zool.* **32**, 1–10 (2016).

58. Hsiung, A. R. et al. Little evidence that lowering the pH of concrete supports greater biodiversity on tropical and temperate seawalls. *Mar. Ecol. Prog. Ser.* **656**, 193–205 (2020).

59. Kaehler, S. & Williams, G. A. Early development of algal assemblages under different regimes of physical and biotic factors on a seasonal tropical rocky shore. *Mar. Ecol. Prog. Ser.* **172**, 61–71 (1998).

60. Williams, G. A., Davies, M. S. & Nagarkar, S. Primary succession on a seasonal tropical rocky shore: the relative roles of spatial heterogeneity and herbivory. *Mar. Ecol. Prog. Ser.* **203**, 81–94 (2000).

61. Tan, S. K. *Land Reclamation in Singapore* (National University of Singapore, Singapore, 1976).

62. Hilton, M. J. & Chou, L. M. Sediment facies of a low-energy, meso-tidal, fringing reef, Singapore. *Singap. J. Trop. Geogr.* **20**, 111–130 (1999).

63. Zhao, K. et al. Modelling surface temperature of granite seawalls in Singapore. *Case Stud. Therm. Eng.* **13**, 100395 (2019).

64. Loke, L. H. L., Bouma, T. J. & Todd, P. A. The effects of manipulating microhabitat size and variability on tropical seawall biodiversity: field and flume experiments. *J. Exp. Mar. Biol. Ecol.* **492**, 113–120 (2017).

65. Strain, E. M. et al. A global analysis of complexity–biodiversity relationships on marine artificial structures. *Glob. Ecol. Biogeogr.* **30**, 140–153 (2021).

66. R Core Team. *R: A Language and Environment for Statistical Computing* (R Foundation for Statistical Computing, 2022); https://www.r-project.org/.

67. Schoener, T. W. Competition and the form of habitat shift. *Theor. Popul. Biol.* **6**, 265–307 (1974).

68. Chisholm, R. A. & Pacala, S. W. Niche and neutral models predict asymptotically equivalent species abundance distributions in high-

diversity ecological communities. *Proc. Natl Acad. Sci. USA* **107**, 15821–15825 (2010).

69. Chisholm, R. A. & Pacala, S. W. Theory predicts a rapid transition from niche-structured to neutral biodiversity patterns across a speciation-rate gradient. *Theor. Ecol.* **4**, 195–200 (2011).

70. Abramowitz, M. & Stegun, I. A. *Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables* (Dover, 1972).

Acknowledgements

We thank E. Heery and N. Kristensen for helpful comments on the manuscript and H. Muller-Landau for helpful discussion. L.H.L.L. was supported by a Macquarie University Research Fellowship (no. 110042722) and a grant from Wildlife Reserves Singapore Conservation Fund. R.A.C. was supported by a grant from the James S. McDonnell Foundation (no. 220020470).

Author information

Authors and Affiliations

1. School of Natural Sciences, Faculty of Science and Engineering, Macquarie University, North Ryde, New South Wales, Australia

   Lynette H. L. Loke

2. Department of Biological Sciences, National University of Singapore, Singapore, Singapore

   Ryan A. Chisholm

Contributions

L.H.L.L. conceived and conducted the experiment, collected the data, performed analyses and wrote the first draft of the manuscript. R.A.C.

contributed ideas, performed analyses and edited the manuscript. All authors contributed substantially to revisions.

Corresponding author

Correspondence to Lynette H. L. Loke.

Ethics declarations

Competing interests

The authors declare no competing interests.

Peer review

Peer review information

Nature thanks David Alonso and the other, anonymous, reviewers(s) for their contribution to the peer review of this work. Peer reviewer reports are available.

Additional information

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Extended data figures and tables

Extended Data Fig. 1 Trends in species richness in the experimental communities over a year.

Each plot shows species richness (\(S\)) against time (months) for one experimental treatment (panels), where treatments vary in the level of immigration (governed by number of holes, \(h\)). Experimental communities stabilized after 6–7 months for all but the highest immigration

treatments. The horizontal and vertical coordinates of points have been jittered to improve visibility. Solid curves are mean estimates with shaded 95% confidence intervals; vertical dashed lines indicate the seventh month from the start of the experiment.

Source Data

Extended Data Fig. 2 Fitting a variant of the classic model.

Each small grey point indicates a single experimental unit (seawall tile; 12 h-treatments, n = 5) in a given month (x-coordinates are jittered to improve visibility); each large black point shows a treatment mean. Fitted models are shown by the solid curves (mean estimates) with shaded 95% confidence intervals from 1,000 bootstraps; the novel model (Eq. 6) is shown in red while the variant of the classic model is in green (Eq. 5; see Methods). This variant of the classic model was substantially poorer than the novel model although better than the original classic model (judged by AIC). The dashed green horizontal line indicates the expected species richness at high levels of immigration under the variant of the classic model, as estimated from the fitted parameters.

Source Data

Supplementary information

Supplementary Information

Supplementary Text, Figs. 1–3 and Table 1.

Reporting Summary

Peer Review File

Source data

Source Data Fig. 3

[Source Data Fig. 4]

[Source Data Extended Data Fig. 1]

[Source Data Extended Data Fig. 2]

Rights and permissions

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by/4.0/.

Reprints and Permissions

About this article

Cite this article

Loke, L.H.L., Chisholm, R.A. Unveiling the transition from niche to dispersal assembly in ecology. *Nature* **618**, 537–542 (2023). https://doi.org/10.1038/s41586-023-06161-x

- Received: 20 September 2022
- Accepted: 02 May 2023
- Published: 07 June 2023

- Issue Date: 15 June 2023
- DOI: https://doi.org/10.1038/s41586-023-06161-x

Share this article

Anyone you share the following link with will be able to read this content:

Get shareable link

Sorry, a shareable link is not currently available for this article.

Copy to clipboard

Provided by the Springer Nature SharedIt content-sharing initiative

What maintains biodiversity in ecological communities?

Nature Research Briefing 07 Jun 2023

This article was downloaded by **calibre** from https://www.nature.com/articles/s41586-023-06161-x

- Article
- Open Access
- Published: 24 May 2023

A median fin derived from the lateral plate mesoderm and the origin of paired fins

- Keh-Weei Tzung,
- Robert L. Lalonde,
- Karin D. Prummel,
- Harsha Mahabaleshwar,
- Hannah R. Moran,
- Jan Stundl,
- Amanda N. Cass,
- Yao Le,
- Robert Lea,
- Karel Dorey,
- Monika J. Tomecka,
- Changqing Zhang,
- Eline C. Brombacher,
- William T. White,
- Henry H. Roehl,
- Frank J. Tulenko,
- Christoph Winkler,
- Peter D. Currie,
- Enrique Amaya,
- Marcus C. Davis,
- Marianne E. Bronner,
- Christian Mosimann &
- …
- Tom J. Carney

Show authors

Nature **volume 618**, pages 543–549 (2023)

- 9499 Accesses
- 168 Altmetric
- Metrics details

Abstract

The development of paired appendages was a key innovation during evolution and facilitated the aquatic to terrestrial transition of vertebrates. Largely derived from the lateral plate mesoderm (LPM), one hypothesis for the evolution of paired fins invokes derivation from unpaired median fins via a pair of lateral fin folds located between pectoral and pelvic fin territories[1]. Whilst unpaired and paired fins exhibit similar structural and molecular characteristics, no definitive evidence exists for paired lateral fin folds in larvae or adults of any extant or extinct species. As unpaired fin core components are regarded as exclusively derived from paraxial mesoderm, any transition presumes both co-option of a fin developmental programme to the LPM and bilateral duplication[2]. Here, we identify that the larval zebrafish unpaired pre-anal fin fold (PAFF) is derived from the LPM and thus may represent a developmental intermediate between median and paired fins. We trace the contribution of LPM to the PAFF in both cyclostomes and gnathostomes, supporting the notion that this is an ancient trait of vertebrates. Finally, we observe that the PAFF can be bifurcated by increasing bone morphogenetic protein signalling, generating LPM-derived paired fin folds. Our work provides evidence that lateral fin folds may have existed as embryonic anlage for elaboration to paired fins.

Main

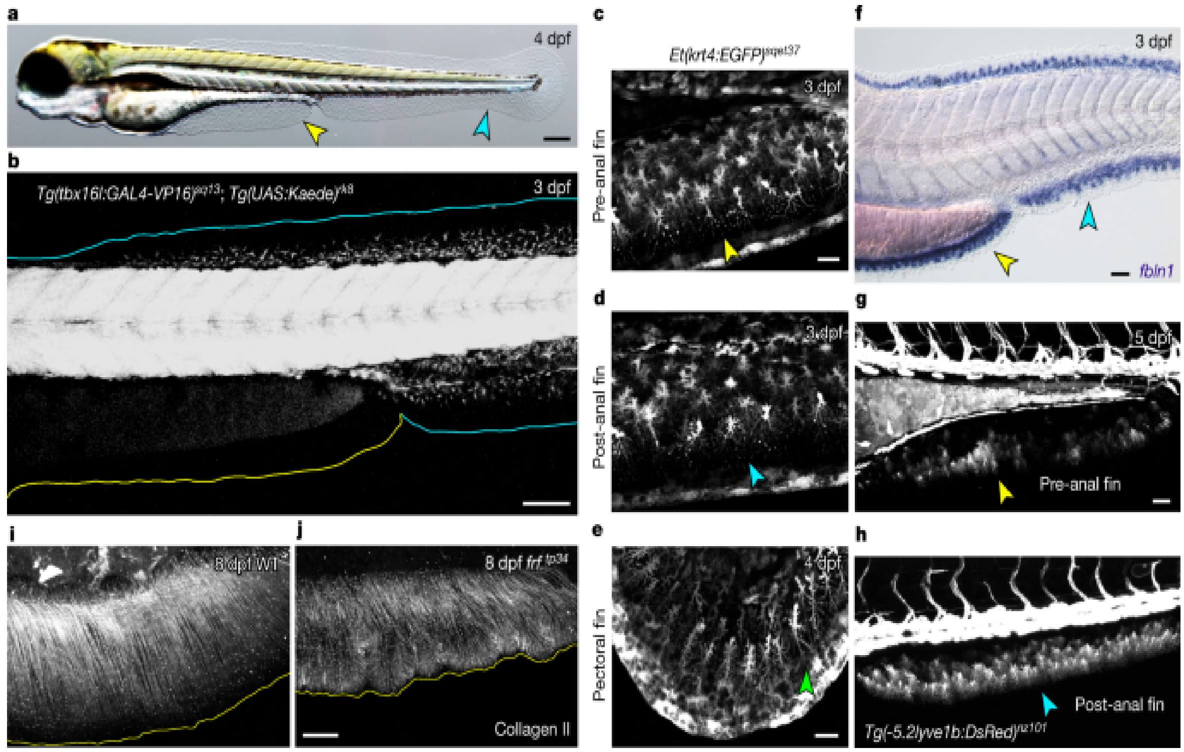
Two alternate hypotheses have been proposed to explain the evolutionary origin of vertebrate paired appendages (fins and limbs). Derivation from posterior gill arches was posited by Gegenbaur[3], whilst a number of

anatomists later invoked a rival theory, the lateral fin fold hypothesis. This proposed that paired fins derived (either phylogenetically or ontogenetically) from longitudinal bilateral fin folds that were then subdivided[4,5]. Whilst recent molecular studies have provided some evidence in support of each hypothesis[2,6,7], there remains significant criticism of the lack of substantiation in the fossil record or in embryology[1,8]. Certain stem vertebrates, including anaspid-related fossils, show evidence of lateral fin folds; however, these fin folds mostly consist of soft tissue with only sporadic skeletal elements and are thus poorly preserved. This has led to conflicting interpretations[9,10,11,12,13]. The developmental programme for paired fins has been postulated to have been first assembled in median fins, which appear in the fossil record before the origin of paired fins[2]. A number of studies have traced the cellular origin of median fins in lamprey, catshark, zebrafish and *Xenopus*[2,14,15]. All median fins, both larval and adult, assessed so far have shown derivation from the paraxial mesoderm (PM), whilst paired fins are known to be derived from the lateral plate mesoderm (LPM). Thus, the median fin programme was most likely transferred to the LPM from the PM[2], possibly before the formation of hypothesized lateral fin folds. How or when this transition occurred is unclear. As only a subset of median fins in zebrafish has been assayed, we expanded the characterization of the composition and origin of median fins in zebrafish to determine if PM derivation was an invariant characteristic.

As with most surveyed jawed vertebrates (gnathostomes), larval zebrafish possess two median unpaired fin folds. A caudal fin fold (or major lobe) runs continuously from the dorsal midline around the caudal end of the larva and then ventrally to the anus (Fig. 1a). A pre-anal fin fold (PAFF; or minor lobe) runs along the underside of the yolk sac extension, immediately anterior to the anus[16] (Fig. 1a). The median fin folds are resorbed during metamorphosis, and the caudal fin fold is replaced by three separate PM-derived median adult fins[14,17]: the caudal fin, the dorsal fin and the anal fin. No adult median fin replaces the PAFF, which is a developmentally transient structure[16]. The mesenchyme of the caudal fin fold can be labelled by the photoconvertible Kaede green-to-red fluorescent protein expressed under control of the *tbx16l* promoter, confirming that this median fin is derived from the PM[18]. In contrast, we consistently failed to observe any Kaede labelling in the PAFF (Fig. 1b and Extended Data Fig. 1a). This labelling

difference is not due to de novo activity of the *tbx16l* promoter specifically in the caudal median fin fold, as we photoconverted Kaede in anterior somites above the PAFF at 24 h post-fertilization (hpf) and traced the photoconverted Kaede to the mesenchyme of the dorsal portion of the caudal fin fold; again, we never found labelled cells in the PAFF ($n = 11$) (Extended Data Fig. 1b–d). The absence of PAFF labelling by a PM-lineage trace is not due to absence of mesenchyme cells, as we observed an abundance of these cells by Nomarski optics at 3 days post-fertilization (dpf) and in the enhancer trap line *Et(krt4:EGFP)^{sqet37}* (*ET37*), which labels all fin mesenchyme cells (Extended Data Fig. 1e,f). The *ET37* line further allowed us to visualize the morphology of the mesenchyme cells, which showed a distally polarized stellate shape, indistinguishable from the mesenchyme of all other fins (Fig. 1c–e and Extended Data Fig. 1f–h). PAFF mesenchymal cells expressed known differentiation markers of fin mesenchyme, including *fibulin1* (*fbln1*) and *integrin beta 3b* (*itgb3b*), and were weakly positive for *lyve1b* reporter activity in the *lyve1b:DsRed2* transgenic line, which also expresses in the post-anal ventral fin mesenchyme[19] (Fig. 1f–h and Extended Data Fig. 1i). Hence, despite their divergent developmental origin, PAFF mesenchyme has a similar morphology and expression profile to fin mesenchyme of the major median fin lobe.

Fig. 1: A non-PM-derived median fin fold.



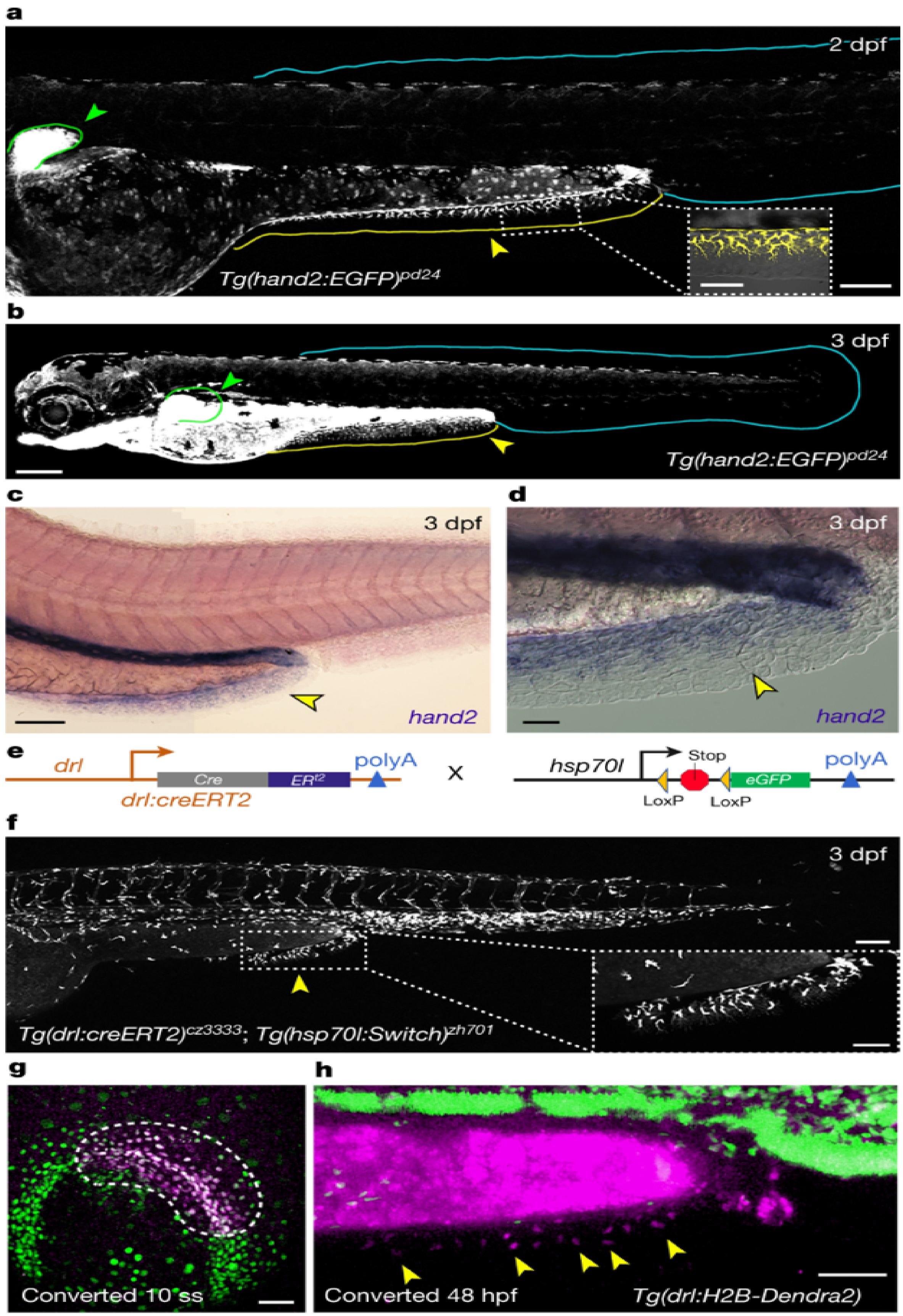
a, Larval 4 dpf zebrafish possess a median PAFF (yellow arrowhead) in addition to the caudal median fin fold (cyan arrowhead). **b**, Confocal image of a 3 dpf *Tg(tbx16l:GAL4-VP16); Tg(UAS:Kaede)* embryo with PM labelled by Kaede showing PM-derived mesenchyme in the caudal median fin fold (cyan outline) but not the PAFF (yellow outline). **c**–**e**, Confocal images of pre-anal (**c**), ventral caudal (**d**) and pectoral (**e**) fins of the *ET37* Enhancer Trap transgenic line indicating that PAFF contains morphologically comparable mesenchyme (indicated by arrowheads) to other larval fin folds. **f**, In situ hybridization of the fin mesenchyme marker *fbln1* in both PAFF (yellow arrowhead) and caudal fin fold (cyan arrowhead) at 3 dpf. **g**,**h**, DsRed expression in mesenchyme of both pre-anal (**g**) and caudal (**h**) fin folds of the 5 dpf *Tg*(-5.2*lyve1b:DsRed)* transgenic line. **i**,**j**, Immunostaining for collagen II in 8 dpf *frf* mutants (**j**) shows loss of fibril organization compared with WT (**i**). Scale bars, 200 µm (**a**); 100 µm (**b**); 20 µm (**c**,**e**); 50 µm (**f**,**g**,**j**).

Few functions of fin mesenchyme cells are defined. The zebrafish *frilly fin* (*frf*) mutants display ruffling of the caudal larval fin fold due to mutations in *bone morphogenetic protein 1a* (*bmp1a*), which is expressed in fin mesenchyme[20], including in the PAFF (Extended Data Fig. 1j). The overall

morphology of the PAFF also displays ruffling in *frf* mutants (Fig. 1i,j). Further, immunostaining of collagen II revealed ordered parallel collagen II fibres in all median fins of the wild type (WT), while these fibres were disorganized in both the caudal median fin folds and the PAFF of *frf* mutants (Fig. 1i,j), indicating that Bmp1a in PAFF mesenchyme is also required for maturation of collagen. Collectively, these results indicate that mesenchymal cells of the PAFF show functional overlap with those of the caudal median fin fold yet have a distinct developmental origin.

The transcription factor Hand2 is expressed in several LPM progenitors, including cardiac, pharyngeal, mesothelial and pectoral fin progenitors[21,22]. Upon examining the *TgBAC(hand2:EGFP)* transgenic line that accurately recapitulates endogenous *hand2* gene expression[23], we observed the expected enhanced green fluorescent protein (eGFP) expression in larval pectoral fins at 2–3 dpf (Fig. 2a,b). Although we observed no *TgBAC(hand2:EGFP)* expression in the major unpaired larval fin lobe, the PAFF mesenchyme was strongly eGFP positive, suggesting an LPM origin (Fig. 2a,b and Extended Data Fig. 2a,b). Additionally, in situ hybridization at 3 and 5 dpf indicated that the PAFF mesenchyme, but no other median fin fold, expressed *hand2* (Fig. 2c,d and Extended Data Fig. 2c,d). If the LPM uniquely contributes to the PAFF, then we might expect that the PAFF, but no other median fin fold, would be affected in an LPM mutant. The zebrafish *hand2* mutant *hands off* (*han^{s6}*) exhibits defects in LPM derivatives, including the heart, pectoral fins and mesothelium[21,22]. Consistent with an LPM origin, the fin height of PAFFs in *han^{s6}* mutant larvae was significantly reduced compared with WT at 3 and 5 dpf, whilst the ventral section of the major median fin lobe was unaffected (Extended Data Fig. 3a,b,g,h,m). We observed similar results with *hand2* antisense morpholino (MO)-based knockdown (Extended Data Fig. 3c–f,n). *hand2* knockdown in the *ET37* line severely reduced the number of eGFP-positive mesenchymal cells in the PAFF at 3 and 5 dpf compared with WT, whilst the mesenchyme of the ventral caudal median fin fold was unaffected (Extended Data Fig. 3i–l,o). Transplantation of *hand2* MO; *ET37* mutant cells into WT suggested that the loss of Hand2 reduced clone size and altered morphology and migration of mesenchyme cells autonomously (Extended Data Fig. 4). Together, these observations are consistent with an LPM origin of PAFF mesenchyme.

Fig. 2: The PAFF is an LPM-derived median fin fold.


a
2 dpf
Tg(hand2:EGFP)pd24
b
3 dpf
Tg(hand2:EGFP)pd24
c
3 dpf
hand2
d
3 dpf
hand2
e
drl
drl:creERT2
Cre
ERt2
polyA
×
hsp70l
Stop
LoxP
LoxP
eGFP
polyA
f
3 dpf
Tg(drl:creERT2)cz3333; Tg(hsp70l:Switch)zh701
g
Converted 10 ss
h
Converted 48 hpf
Tg(drl:H2B-Dendra2)

a,**b**, Confocal images of *Tg(hand2:EGFP)* embryos at 2 dpf (**a**) and 3 dpf (**b**) showing eGFP labelling of the mesenchyme of the pectoral (green outline and arrowheads) and PAFFs (yellow outline and arrowheads, and magnified in inset) but not the caudal fin fold (cyan outline). **c**,**d**, In situ hybridization of *hand2* at 3 dpf shows fin expression of *hand2* only in the PAFF (yellow arrowheads (**c**), and higher magnification with Nomarski optics indicates expression in the mesenchyme (**d**). **e**, Schematic of the LPM lineage tracing transgenes. **f**, Lineage tracing of LPM using transgenics depicted in (**e**) following 4-OHT treatment and heat shock before imaging shows that PAFF mesenchyme is derived from the LPM (yellow arrowhead and magnified in inset). **g**,**h**, Ventral (**g**) and lateral (**h**) confocal images of the *drl:H2B-Dendra2* transgenic line at the 10-somite stage (10 ss) (**g**) and 48 hpf (**h**) following ultra-violet laser photoconversion in the region of the LPM outlined in (**g**). **h**, Photoconverted PAFF mesenchyme is indicated by yellow arrowheads. Scale bars, 100 µm (**a**,**c**,**f**); 50 µm (**a** (inset),**f** (inset),**g**,**h**); 200 µm (**b**); 20 µm (**d**).

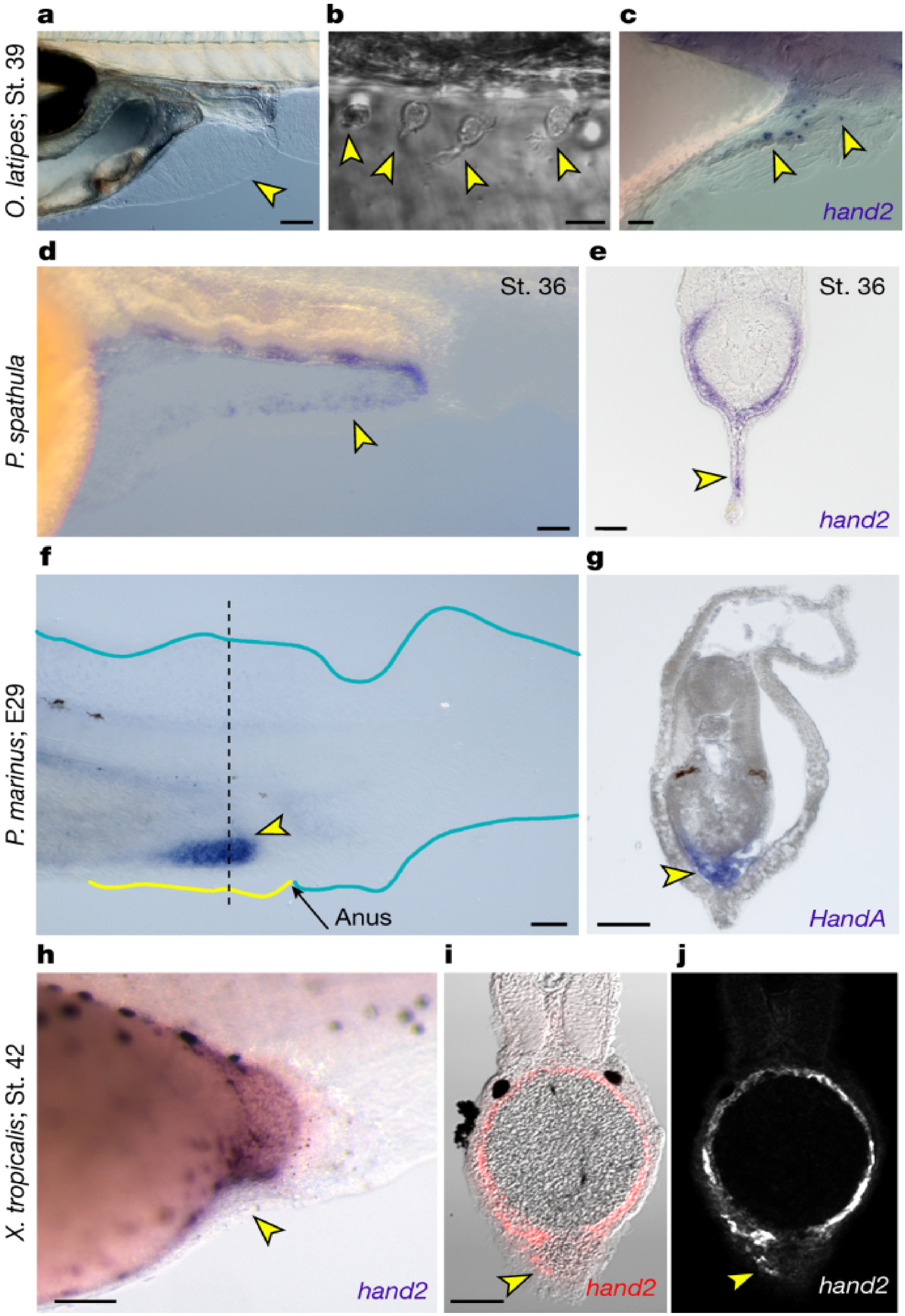
To confirm that this expression of *hand2* indicates an LPM origin of the PAFF rather than de novo expression, we permanently labelled the LPM in a mosaic fashion using a *Tg(drl:creERT2; hsp70l:Switch)* transgenic combination[24] (Fig. 2e). During gastrulation and early somitogenesis, *drl:creERT2* is expressed in LPM-primed mesoderm with increasing selectivity to the LPM, suitable for lineage labelling of this mesodermal lineage using 4-OH-tamoxifen (4-OHT) induction and *loxP* reporters[24]. Consistent with an LPM origin of paired appendages, we first documented that this LPM lineage tracing approach labels cells of the paired fins, in particular the larval pectoral fins and pelvic fins. We found lineage labelling in the endoskeletal disc and mesenchyme of 4.5 dpf pectoral fins and also, the intraray fibroblasts and Zns5-positive osteoblasts of the adult pelvic fins (Extended Data Fig. 5a–d). In addition, we observed LPM lineage labelling of the PAFF mesenchyme at 3 and 5 dpf, whilst the labelling in the caudal fin fold was limited to myeloid cells, which are also LPM derived (Fig. 2f and Extended Data Fig. 5e–g). These data further support the contribution of the LPM to the PAFF mesenchyme but not to other median fin folds. We next used *drl*-driven expression of the Dendra2 green-to-red photoconvertible fluorophore to localize the precise LPM origin of PAFF mesenchyme. During early to mid-somitogenesis (8–10-somite stage (ss)),

drl:H2B-Dendra2 is expressed in the nuclei of the LPM (Extended Data Fig. 6a,b). Photoconversion of posterior-most LPM at this stage, using ultra-violet laser illumination, resulted in selective Dendra2-red labelling of this field of LPM (Fig. 2g and Extended Data Fig. 6c,d). These cells were then traced to the migrating PAFF mesenchyme at 40 and 48 hpf ($n = 4$) (Fig. 2h and Extended Data Fig. 6e–j,q–t), whilst unconverted control cells did not show any Dendra2-red PAFF labelling (Extended Data Fig. 6k–p,u–x). This confirmed PAFF mesenchyme originates from the LPM, with significant contribution from the posterior-most LPM at early segmentation stages.

To investigate whether an LPM-seeded median PAFF is a zebrafish apomorphy or if it is a more ancestral feature shared with other vertebrates, we examined the larvae of other taxa for the presence of a PAFF and used *hand2* in situ hybridization as a proxy marker of an LPM origin of constituent mesenchymal cells. Medaka (*Oryzias latipes*) larvae have a markedly smaller PAFF that possesses only a sparse number of mesenchymal cells as detected by Nomarski imaging (Fig. 3a,b). These mesenchymal cells have cell extensions projecting distally and were mostly proximal in the fin where the interstitial space is thickest (Fig. 3b). Accordingly, medaka *hand2* was expressed proximally in a punctate pattern in PAFFs, consistent with mesenchymal expression, and was not expressed in other median fin folds (Fig. 3c and Extended Data Fig. 7a–c). To confirm this, we expressed eGFP in early medaka LPM by injection of a -*6.35drl:EGFP* construct. We saw expression in mesenchymal cells in the PAFF transiently and also, in stable transgenic embryos (eight PAFF mesenchyme cells seen in four stage 36 *Tg(-6.35drl:EGFP)* transgenic medaka embryos) (Extended Data Fig. 7d,e). In addition, we used 1,1′-dioctadecyl-3,3,3′,3′-tetramethylindocarbocyanine perchlorate (DiI) lipophilic dye labelling to label the posterior LPM territory at stage 20 (four-somite stage). Consistent with our positional mapping of PAFF-seeding mesenchyme in zebrafish, posterior LPM injection of DiI consistently labelled mesenchyme in the medaka PAFF (total of 27 cells in nine of nine labelled embryos) (Extended Data Fig. 7f–i). A basal actinopterygian, the American paddlefish (*Polyodon spathula*), also forms a transient larval PAFF[8]; we identified the expression of *P. spathula hand2* in PAFF mesenchymal cells between stages 36 and 39, but *hand2* expression was notably absent from the caudal median fin fold (Fig. 3d,e and Extended Data

Fig. 7j–l). PAFF *hand2* mesenchyme expression persisted until PAFF regression (around stages 45 and 46). From stage 39, we observed *hand2* expression in the core of the nascent pelvic fins (Extended Data Fig. 7k,l). We conclude that a larval PAFF with an LPM mesenchyme is ancestral for actinopterygians. To establish if an LPM contribution to median fins occurred before the origin of paired fins, we examined the larvae of a jawless vertebrate, the sea lamprey (*Petromyzon marinus*). Lampreys also possess a small transient larval PAFF[25], within which strong, specific expression of the lamprey *hand2* orthologue, *HandA,* was detected (Fig. 3f,g), although there was evidence of fainter *HandA* expression in the dorsal anterior fin. Finally, both the embryos of the sarcopterygian lungfish[26] and amphibian tadpoles (e.g., *Xenopus laevis*[27]*, Xenopus tropicalis* and axolotls[15]) also possess a small transient PAFF (Fig. 3h). Fluorescence in situ hybridization demonstrated individual *hand2*-positive mesenchymal cells invading this fin in *X. tropicalis* but not the major median fin lobe at stage 42 (Fig. 3h–j). We confirmed this LPM contribution by injecting -*6.35drl:EGFP* into *X. laevis* embryos. *X. laevis* has a larger PAFF than *X. tropicalis,* and we observed extensive transient eGFP labelling in the PAFF compared with the caudal fin fold (Extended Data Fig. 7m–o). Thus, we show an LPM origin for PAFF mesenchyme in the larvae of representatives of major extant vertebrate linages: cyclostomes, actinopterygians and sarcopterygians.

Fig. 3: PAFF mesenchyme expression of *hand2* is conserved across vertebrates.



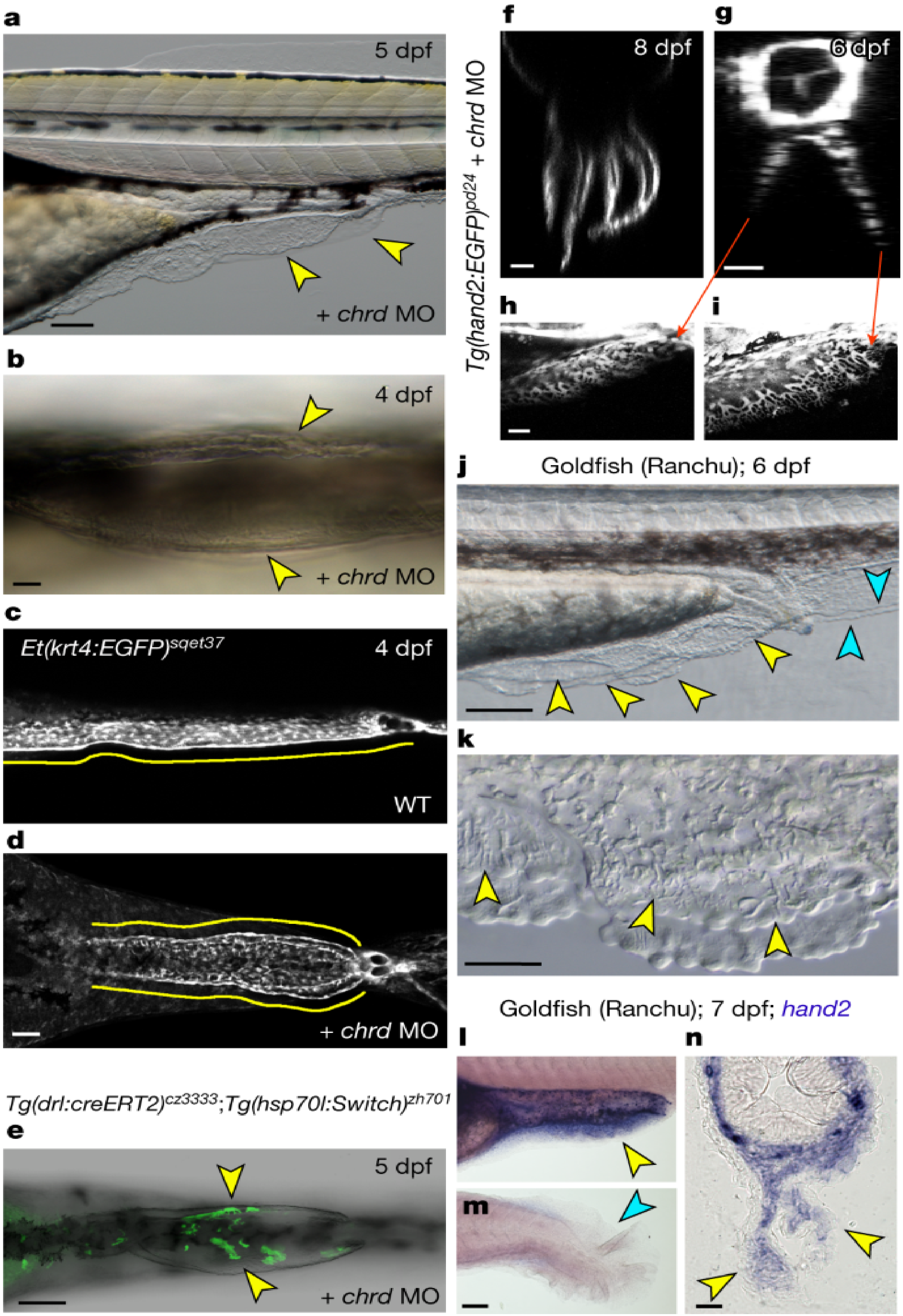
a
b
c
O. latipes; St. 39
hand2
d
e
P. spathula
St. 36
St. 36
hand2
f
g
P. marinus; E29
Anus
HandA
h
i
j
X. tropicalis; St. 42
hand2
hand2
hand2

a,**b**, Nomarski images of 9 dpf medaka showing PAFF (**a**) with dispersed mesenchymal cells (yellow arrowheads) (**b**). **c**, In situ hybridization of medaka at stage 39 showing *hand2* expression in pre-anal fin mesenchyme (yellow arrowheads). **d**,**e**, In situ hybridization of *hand2* in stage 36 paddlefish embryos shown laterally (**d**) or in transverse section (**e**). *hand2*-positive PAFF fin mesenchyme is indicated by yellow arrowheads. **f**,**g**, *HandA* in situ hybridization of stage E29 lamprey (*P. marinus*) embryos shown laterally (**f**) or in transverse section (**g**). Lamprey show strong expression of *HandA* in a fin anterior to the anus (yellow arrowhead in **f**) corresponding to cells in the interior of the fin (**g**). The section location of (**g**) is indicated by the dashed line in (**f**). **h**–**j**, Chromogenic (**h**) or fluorescent (**i**,**j**) in situ hybridization of *hand2* in stage 42 *X. tropicalis* embryos shown laterally (**h**) or in transverse section (**i**,**j**). Fluorescent image (**j**) is overlayed on the Nomarski image (**i**). The small PAFF in *Xenopus* contains sparse *hand2*-positive fin mesenchyme (yellow arrowheads). St., stage. Scale bars, 100 μm (**a**,**d**,**h**); 10 μm (**b**); 20 μm (**c**); 50 μm (**e**–**g**,**i**).

Given the LPM origin of the median PAFF, we considered that it may have importance in the transition from PM-derived median fins to LPM-derived paired fins. We tested if this LPM median fin fold could be duplicated into paired LPM-derived lateral fin folds. Zebrafish larvae mutant for the bone morphogenetic protein (BMP) antagonist *chordin* develop ventral caudal fin fold duplications, although whether the PAFF is duplicated in these mutants has not been reported[28]. Injection of a low dose of *chrd* MO into zebrafish recapitulated various ventralized phenotypes, which included individuals with duplicated or multiple PAFFs (Fig. 4a,b,f and Extended Data Fig. 8a,b). We visualized these PAFF phenotypes by *chrd* MO injection into the *ET37* transgenic line, which demonstrated that each fin fold of duplicated PAFFs contained mesenchyme (Fig. 4c,d). Lineage tracing using the *Tg(drl:creERT2; hsp70l:Switch)* transgenic combination injected with *chrd* MO demonstrated that the mesenchyme cells of the duplicated PAFFs are indeed derived from the LPM (Fig. 4e). By generating *chrd* morphants in the *Tg(hand2:EGFP*) line and then imaging by light-sheet microscopy, we generated three-dimensional reconstructions documenting duplicated PAFFs along the yolk extension, all of which harboured eGFP-positive mesenchymal cells, further demonstrating that the mesenchyme of these duplicate fin folds was LPM derived (Fig. 4g–i, Extended Data Fig. 8c–j and

Supplementary Video 1). Of note is the generation of multiple parallel fin folds in individual embryos following *chrd* MO injection (Fig. 4f). These observations indicate that bifurcation of the LPM-seeded PAFF readily arises from modulating ventral BMP signalling, such as by reducing the dose of the dorsal inhibitor Chordin.

Fig. 4: Duplication of the PAFF into paired fin folds.



a
5 dpf
+ chrd MO
b
4 dpf
+ chrd MO
c
Et(krt4:EGFP)sqet37
4 dpf
WT
d
+ chrd MO
Tg(drl:creERT2)cz3333;Tg(hsp70l:Switch)zh701
e
5 dpf
+ chrd MO
Tg(hand2:EGFP)pd24 + chrd MO
f
8 dpf
g
6 dpf
h
i
j
Goldfish (Ranchu); 6 dpf
k
Goldfish (Ranchu); 7 dpf; hand2
l
m
n

a,**b**, Lateral (**a**) and ventral (**b**) Nomarski images of PAFFs of 5 dpf (**a**) and 4 dpf (**b**) embryos injected with MO targeting *chordin*, resulting in duplication of PAFF (yellow arrowheads). **c**,**d**, Confocal micrographs, imaged ventrally, of PAFFs in 4 dpf *ET37* transgenic larvae uninjected (**c**) or injected with a *chrd* MO (**d**). PAFFs are indicated with yellow lines. **d**, Note the duplicated anal openings following *chrd* MO injection. **e**, Ventral confocal image of *chrd* MO duplicated PAFFs of the transgenic line shown above. Hydroxytamoxifen treatment at the 12-somite stage and heat shock before imaging show that mesenchyme of duplicated PAFFs is derived from the LPM (yellow arrowheads). **f**–**i**, Light-sheet (**f**) and confocal (**g**–**i**) images of *Tg(hand2:EGFP)* larvae at 8 dpf (**f**) and 6 dpf (**g**–**i**). Orthogonal display through the *x*–*z* plane (**f**,**g**) shows that multiple PAFFs can form and that duplicated PAFFs contain eGFP-positive mesenchyme (**g**–**i**). Lateral views of left (**h**) and right (**i**) duplicated PAFFs of sample in **g** are given. **j**,**k**, Lateral low- (**j**) and high-power (**k**) Nomarski images of the PAFF of the Ranchu goldfish at 6 dpf. The multiple PAFFs (**j**) and individual mesenchyme cells (**k**) are indicated by yellow arrowheads. Duplicated caudal fin folds are indicated by cyan arrowheads (**j**). **l**–**n**, Lateral (**l**,**m**) and transverse (**n**) views of pre-anal (**l**,**n**) and caudal (**m**) fin folds of 7 dpf Ranchu larvae stained by in situ hybridization for *hand2*, where *hand2*-positive PAFF mesenchyme is indicated by yellow arrowheads. Absence of *hand2* in the caudal fin fold is indicated by the cyan arrowhead (**m**). Scale bars, 100 µm (**a**,**e**,**m**); 20 µm (**b**,**f**–**h**,**n**); 50 µm (**d**,**k**); 200 µm (**j**).

To determine if the generation of multiple PAFFs could have evolved spontaneously as part of natural variation, we examined the twin-tail goldfish strain, Ranchu, which has been shown to display bifurcated fin folds including PAFFs due to a loss-of-function mutation in a *chordin* paralogue (*chdA*^{E127X})[29,30,31]. The bifurcated PAFFs and caudal fin folds of larval Ranchu appeared similar to those of zebrafish larvae injected with low doses of *chrd* MO (Fig. 4j and Extended Data Fig. 8k,l). These PAFFs were all populated by mesenchymal cells (Fig. 4k). Using *hand2* in situ hybridization as a proxy of LPM lineage to infer the origin of the Ranchu PAFF mesenchyme, we observed that the core of both duplicated PAFFs expressed *hand2*, whilst the caudal fin fold did not (Fig. 4l–n and Extended Data Fig. 8m,n). This suggests that, as in the zebrafish PAFF, the bifurcated Ranchu PAFFs are also LPM derived. Notably, in addition to those

displaying simple PAFF duplication, some individuals had three or more parallel PAFFs (Fig. 4j and Extended Data Fig. 8n). Thus, we conclude that LPM-derived paired fin folds that can be readily generated in zebrafish larvae have also arisen spontaneously in twin-tail goldfish; thus, they represent a viable morphological innovation. Notably, duplicated LPM-derived PAFFs were situated in ventrolateral locations between the future pectoral and pelvic fin domains, overlying the proposed competence stripes of appendage formation in gnathostomes[32].

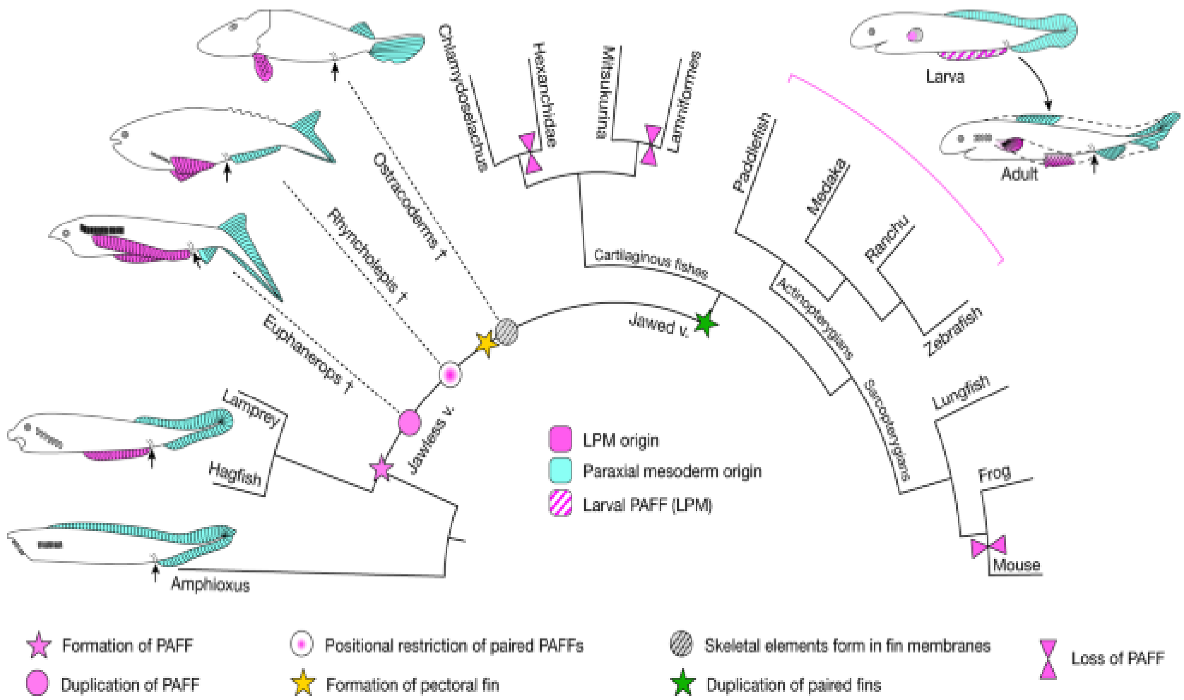
Combining several models and techniques, we here have identified a larval median fin with a mesenchyme core derived from LPM and propose it as an intermediate between PM-derived median fins and LPM-derived paired fins. While apparently conserved from cyclostomes to amphibia, most extant chondrichthyan embryos do not appear to have a discernible PAFF. We used micro-computed tomography (microCT) to examine prehatching stages of the elasmobranch Epaulette shark *Hemiscyllium ocellatum*, which did not show the presence of a clear pre-anal fin (Extended Data Fig. 9a–i). Although this pattern appears to support a loss of the PAFF somewhere in the chondrichthyan lineage, we cannot rule out the possibility that a PAFF arose independently in cyclostomes and Osteichthyes. It is interesting to note that embryos and adults of the frilled shark, *Chlamydoselachus*, do show bifurcated tropeic folds in the ventral midline (Extended Data Fig. 9j–m), structures previously invoked as supporting the fin fold hypothesis[33].

PAFFs are mostly transient larval structures that generally lack a mineralized skeleton, which may explain their poor documentation in the fossil record. Although a true PAFF does not persist into adulthood in most species, it is seen in hagfish adults, where unlike the caudal median fin, it does not possess cartilage rays[34]. In the fossil record, a pre-anal fin is seen in some stem vertebrate fossils, including *Haikouella* and *Haikouichthys*, while the pre-anal fin of *Kerreralepis* shows well-developed plates[35,36]. Zebrafish larval fin mesenchyme is known to persist and contribute to the lepidotrichia of adult fins[18], and LPM-derived dermal fibroblasts form cartilage during axolotl limb regeneration[37]. These data suggest the potential of the PAFF mesenchyme to generate skeletal outcomes.

Spontaneous duplication in twin-tail goldfish species indicates that paired LPM-derived PAFFs could have readily arisen during evolution. In Ranchu, we frequently observed multiple PAFFs, permitting retention of a median PAFF and divergence of duplicates into paired fins. Although there is no evidence of lateral fin folds present in extant vertebrates, adult specimens of the hagfish species *Neomyxine biniplicata* possess anterior lateral paired folds that terminate close to the pre-anal fin[38], resembling partial duplications of the pre-anal fin, although their relevance to paired fin evolution is disputed[39]. In the fossil record, there is evidence that both mineralized and unmineralized lateral paired fin folds were common in anaspids, although homology of anapsid and gnathostome paired fins is highly contentious, as is their relevance to paired fin evolution and the lateral fin fold hypothesis[1,11]. *Pharyngolepis*, *Cowielepis* and *Euphanerops* display pre-anal ventrolateral paired ribbon structures or triangular fins with radials, while in *Jamoytius,* the ribbons lacked mineralized structures[12,40,41]. The galeaspid, *Tujiaaspis*, has been recently described bearing skeletal ventrolateral fins composed of skeletal units[42]. Subsequent regionalization of these adult paired elongate fins through anterior restriction to the pectoral fin region is proposed and exemplified by *Pharyngolepis* and *Rhyncholepis*[11,43].

Our work gives weight to a model for paired fin evolution through co-option of a larval median fin programme to the LPM followed by fin duplication and subsequent regionalization to pectoral and pelvic fins. We present one possible model for how the PAFF may have led to the generation of LPM-derived paired fins (Fig. 5). The PAFF may have originated as a small ectoderm-only fin fold (as seen in amphioxus larvae[44]) with an LPM contribution evolving subsequently upon alterations in LPM topology, such as persistence of a somatopleure and/or a lateral mesodermal divide[45,46]. Similar LPM tissue contexts may have then led to PAFF elongation and following duplication, paired fin regionalization. As the PAFF possesses characteristics of both unpaired and paired fins, the PAFF may represent a novel evolutionary module or at least demonstrate components of the developmental mechanisms that contributed to the emergence of paired appendages.

Fig. 5: Hypothesis of the elaboration of the PAFF to paired fins.



Simplified evolutionary scenario of vertebrates showing the presence of a PAFF and subsequent modifications leading to paired fins. Dashed lines and dagger symbols indicate extinct lineages, and solid lines indicate extant lineages. PM-derived fins and fin folds are in cyan, while LPM-derived fins are in pink. Larval PAFF is hatched. Black arrows indicate the position of the anus.

Methods

Animal husbandry and experiments

Adult zebrafish and medaka were maintained under standard conditions, and embryos were obtained through natural crosses, kept in E3 medium at 28.5 °C and staged according to Iwamatsu[47] and Kimmel et al.[48]. Embryos of twin-tail goldfish (Ranchu strain) were purchased from a local breeder in Singapore, cultured in E3 medium (with methylene blue) at room temperature and staged as per Li et al.[49]. The experimental procedures and protocols were approved by the institutional animal care and usage committees (IACUCs) at the Institute of Molecular and Cell Biology, Agency for Science, Technology and Research, Singapore (IACUC number 140924); Nanyang Technological University (IACUC number A18002);

University of Zurich and the veterinary office of the Canton of Zurich; and the University of Colorado School of Medicine Anschutz Medical Campus (protocol number 979). Paddlefish embryos (*P. spathula*) were obtained and staged as previously described[8], and experiment protocols and animal care were approved by James Madison University (IACUC number 20-1601). Lamprey embryo work was performed based on published protocols and staged per Tahara[50], with experimental protocols approved by the IACUC of the California Institute of Technology (IACUC number 1436). Experiments involving *X. tropicalis* and *X. laevis* were approved by the local ethics committee at the University of Manchester and the Home Office (number PFDA14F2D). *X. laevis* embryos were injected with 100 pg of the *-6.35drl:EGFP* construct[24] together with 400 pg of in vitro-transcribed membrane–*mCherry* messenger RNA in each blastomere at the two-cell stage. Medaka embryos were obtained from the National University of Singapore under IACUC numbers BR19-0120 and BR22-1497. Embryos of Epaulette shark were obtained from brood stock at Monash University as approved by the Monash Animal Ethics Committee under license (number 30347) and staged per Ballard et al.[51].

Fish lines and MOs

The zebrafish WT strain used in this study was AB, while the following mutants were used: *frf*^{tp34} (ref. [20]) and *han*^{S6} (ref. [22]). The transgenic lines used were *ET37*, *Tg(lyve1b:DsRed2)*^{nz101}, *Tg(tbx16l:GAL4-VP16)*^{sq13}*; Tg(UAS:Kaede)*^{rk8}, *Tg*(*hand2:EGFP*)^{pd24} and *Tg*(*drl:creERT2)*^{cz3333}; *Tg(hsp70l:switch*)^{zh701} as described previously[18,23,24,52]. The *drl:H2B-Dendra2* transgenic line was generated by *Tol2*-mediated transgenesis using a plasmid construct assembled through the Multisite Gateway system with LR Clonase II Plus (Life Technologies) and components from the Tol2kit[53]. The medaka used was the Cab strain, and the *-6.35drl:EGFP* transgenic line was generated by *I-SceI* meganuclease-mediated transgenesis[54] using a plasmid that contained the *-6.35drl:EGFP* sequence[24] flanked by two I-SceI recognition sites. The *hand2* MO sequence was (5′–3′) CCTCCAACTAAACTCATGGCGACAG and injected at 500 µM. The *chordin* MO sequence was (5′–3′) ATCCACAGCAGCCCCTCCATCATCC and injected at 200 µM.

In situ hybridization

In situ hybridization on zebrafish, medaka and goldfish embryos was conducted as described[55]. Probes used were *fbln1* (ref. [56]), *itgb3b*[57], *bmp1a*[20], zebrafish *hand2* (ref. [22]) and medaka *hand2* (ref. [58]). Goldfish *hand2* probe was generated by polymerase chain reaction using goldfish complementary DNA and the primer sequences (5′–3′) ACGTTTTATGGGGAGACAACC and <u>TAATACGACTCACTATAGGG</u>TCTTCCTTGGCGTCTGTCTT, where the T7 RNA polymerase site is underlined. The transverse section of whole-mount in situ hybridization was achieved via cryosections. Briefly, stained embryos were embedded in agar sucrose and cryoprotected in 30% sucrose. Samples were embedded in OCT medium (Tissue-Tek) and cut into 16- to 20-µm sections in a Leica cryostat (CM3050S). Paddlefish in situ hybridization was performed as previously stated, and the paddlefish *hand2* probe was used[59]. Whole-mount chromogenic in situ hybridizations and fluorescent in situ hybridizations and sections were performed as previously described[60,61] on *X. tropicalis* larvae using a *hand2* antisense probe from the expressed sequence tag (EST) clone (TTba012k13) identified from the *X. tropicalis* full-length EST library[62]. Lamprey in situ hybridization was accomplished using *HandA* as a probe according to the published protocol[63].

Transplantation

The *hand2* cell autonomy experiments were performed by transplanting marginal cells at gastrula stages from *ET37* hosts to the margin of unlabelled WT hosts as previously described[64]. Donors were injected with 50 pg *H2B–mCherry* messenger RNA either with or without 500 µM *hand2* MO. Chimeras were then raised, and PAFF mesenchyme was imaged at 3 dpf.

DiI injection of medaka

CellTracker CM-DiI Dye (C7000; Invitrogen) was microinjected using filamented glass capillaries at a concentration of 1 $\mu g\ \mu l^{-1}$ into the LPM, just posterior to Kupffer's vesicle at the approximately four-somites stage using Eppendorf FemtoJet.

Immunostaining

Antibody staining was executed as previously described[20] using either zebrafish embryos or sections. The primary antibodies, sources and dilutions used were anti-Col2A1 (II-II6B3; DSHB; 1:100), anti-eGFP (TP401; Torrey Pines; 1:500), zns-5 (AB_10013796; ZIRC; 1:200) and anti-SM22 alpha/Transgelin (ab14106; Abcam; 1:250). Secondary antibodies were all from Invitrogen and used at 1:500: Alexa 488-conjugated donkey anti-rabbit immunoglobulin G (IgG; catalogue number A-21206), Alexa 546-conjugated donkey anti-mouse IgG (catalogue number A10036), Alexa 647-conjugated donkey anti-rabbit IgG (catalogue number A-31573) and Alexa 647-conjugated donkey anti-mouse IgG (catalogue number A-31571).

Microscopy

Confocal images were taken on an LSM800, LSM880 or LSM980 Zeiss microscope. For high magnification of bright-field or Nomarski images, a Zeiss AxioImager M2 was used. A Zeiss AxioZoom V16 was utilized to capture low magnification of bright-field images. Light-sheet fluorescence microscopy was undertaken using a Zeiss Lightsheet Z.1, and sample preparation and experimental procedures were performed based on the manufacturer's manual. Microscopy images (fluorescence and bright field) were analysed by ZEN Blue v.3.6 (Zeiss), Fiji (ImageJ v.1.52p) and IMARIS v.9.9.1 (Oxford Instruments).

High-resolution X-ray computed tomography

Shark samples previously fixed in 4% paraformaldehyde (PFA) and dehydrated for freezer storage were stained with 1% I_2 dissolved in ethanol according to Metscher[65]. X-ray computed tomography scans were done using a Zeiss Xradia 520 Versa system. Images were generated using Avizo 3D software (ThermoFisher Scientific).

Photoconversion

To track cell lineage, Kaede and Dendra-2 photoconversion was conducted as previously described[18]. In brief, 24 hpf Kaede or Dendra-2 green-expressing embryos were mounted in 3% methyl cellulose or 0.5% low-melting point agarose and imaged by confocal microscopy. A selected region of interest was photoconverted to express red fluorescence using 405-nm ultra-violet laser illumination. Embryos were imaged immediately after photoconversion by sequential acquisition using 488- and 561-nm-wavelength channels to confirm successful conversion. Then, the same embryos were re-imaged at 40, 48 or 72 hpf.

Tamoxifen treatment

To induce Cre activity in *Tg*(*drl:creERT2, hsp70l:Switch*) transgenic embryos, 4-OHT was added in E3 (final concentration of 10 μM) at the 12-somite stage. Then, 4-OHT was washed off at 24 hpf, and embryos were raised until the desired stage. For specific embryonic and larval time points, embryos were heat shocked for 1 h at 37 °C to initiate the expression of eGFP (2–3 h before fixation). Subsequently, embryos were sampled, fixed and stained with 4′,6-diamidino-2-phenylindole (DAPI), and confocal imaging was undertaken.

Adult zebrafish were heat shocked for 16 h at 37 °C, anesthetized with Tricaine and euthanized in an ice-water bath. Successfully switched individuals were selected based on eGFP-expressing heart (major derivate organ of the *drl*-expressing LPM), and the paired pectoral and pelvic fins and anal fin were dissected and fixed overnight with 4% PFA.

Statistical analysis and reproducibility

Statistical analysis was performed using Prism v.9 (GraphPad). A two-tailed Mann–Whitney test was used when two conditions were compared. No statistical methods were used to predetermine sample size. All experiments were performed at least three times on different weeks with different biological samples. The experiments were not randomized, and investigators were not blinded. Consistent labelling within each batch of embryos was confirmed, and representative samples were used for imaging.

Reporting summary

Further information on research design is available in the Nature Portfolio Reporting Summary linked to this article.

Data Availability

The authors confirm that all relevant data are provided in this paper and in its Extended Data files. The data for measurements of fin size and cell number in Extended Data Fig. 3m–o are available in Figshare with the identifier https://doi.org/10.6084/m9.figshare.22269769. Source data are provided with this paper.

References

1. Coates, M. I. The evolution of paired fins. *Theory Biosci.* **122**, 266–287 (2003).

2. Freitas, R., Zhang, G. & Cohn, M. J. Evidence that mechanisms of fin development evolved in the midline of early vertebrates. *Nature* **442**, 1033–1037 (2006).

3. Gegenbaur, C. *Elements of Comparative Anatomy* (Macmillan and Company, 1878).

4. Balfour, F. M. On the development of the skeleton of the paired fins of Elasmobranchii, considered in relation to its bearings on the nature of the limbs of the Vertebrata. *Proc. Zool. Soc. Lond.* **49**, 656–670 (1881).

5. Mivart, S. G. Notes on the fins of elasmobranchs, with considerations on the nature and homologues of vertebrate limbs. *Trans. Zool. Soc. Lond.* **10**, 439–484 (1879).

6. Gillis, J. A., Dahn, R. D. & Shubin, N. H. Shared developmental mechanisms pattern the vertebrate gill arch and paired fin skeletons. *Proc. Natl Acad. Sci. USA* **106**, 5720–5724 (2009).

7. Diogo, R. Cranial or postcranial—dual origin of the pectoral appendage of vertebrates combining the fin-fold and gill-arch theories? *Dev. Dyn.* **249**, 1182–1200 (2020).

8. Bemis, W. E. & Grande, L. in *Mesozoic Fishes 2: Systematics and Fossil Record*. (eds Arratia, G. & Schultze, H.-P.) 41–68 Vrelag Dr. F. Pfeil, 1999).

9. Zhang, X. G. & Hou, X. G. Evidence for a single median fin-fold and tail in the Lower Cambrian vertebrate, *Haikouichthys ercaicunensis*. *J. Evol. Biol.* **17**, 1162–1166 (2004).

10. Shu, D. G. et al. Lower Cambrian vertebrates from south China. *Nature* **402**, 42–46 (1999).

11. Wilson, M. V. H., Hanke, G. F. & Märss, T. in *Major Transitions in Vertebrate Evolution* (eds Anderson, J. & Sues, H.-D.) Ch. 3, 122–149 (Indiana Univ. Press, 2007).

12. Sansom, R. S., Freedman, K. I. M., Gabbott, S. E., Aldridge, R. J. & Purnell, M. A. Taphonomy and affinity of an enigmatic Silurian vertebrate, *Jamoytius kerwoodi* White. *Palaeontology* **53**, 1393–1409 (2010).

13. Sansom, R. S., Gabbott, S. E. & Purnell, M. A. Unusual anal fin in a Devonian jawless vertebrate reveals complex origins of paired appendages. *Biol. Lett.* **9**, 20130002 (2013).

14. Lee, R. T., Thiery, J. P. & Carney, T. J. Dermal fin rays and scales derive from mesoderm, not neural crest. *Curr. Biol.* **23**, R336–R337 (2013).

15. Taniguchi, Y. et al. Mesodermal origin of median fin mesenchyme and tail muscle in amphibian larvae. *Sci. Rep.* **5**, 11428 (2015).

16. Parichy, D. M., Elizondo, M. R., Mills, M. G., Gordon, T. N. & Engeszer, R. E. Normal table of postembryonic zebrafish development: staging by externally visible anatomy of the living fish. *Dev. Dyn.* **238**, 2975–3015 (2009).

17. Shimada, A. et al. Trunk exoskeleton in teleosts is mesodermal in origin. *Nat. Commun.* **4**, 1639 (2013).

18. Lee, R. T., Knapik, E. W., Thiery, J. P. & Carney, T. J. An exclusively mesodermal origin of fin mesenchyme demonstrates that zebrafish trunk neural crest does not generate ectomesenchyme. *Development* **140**, 2923–2932 (2013).

19. Okuda, K. S. et al. A zebrafish model of inflammatory lymphangiogenesis. *Biol. Open* **4**, 1270–1280 (2015).

20. Asharani, P. V. et al. Attenuated BMP1 function compromises osteogenesis, leading to bone fragility in humans and zebrafish. *Am. J. Hum. Genet.* **90**, 661–674 (2012).

21. Prummel, K. D. et al. Hand2 delineates mesothelium progenitors and is reactivated in mesothelioma. *Nat. Commun.* **13**, 1677 (2022).

22. Yelon, D. et al. The bHLH transcription factor hand2 plays parallel roles in zebrafish heart and pectoral fin development. *Development* **127**, 2573–2582 (2000).

23. Yin, C., Kikuchi, K., Hochgreb, T., Poss, K. D. & Stainier, D. Y. R. Hand2 regulates extracellular matrix remodeling essential for gut-looping morphogenesis in zebrafish. *Dev. Cell* **18**, 973–984 (2010).

24. Mosimann, C. et al. Chamber identity programs drive early functional partitioning of the heart. *Nat. Commun.* **6**, 8146 (2015).

25. Richardson, M. K. & Wright, G. M. Developmental transformations in a normal series of embryos of the sea lamprey *Petromyzon marinus* (Linnaeus). *J. Morphol.* **257**, 348–363 (2003).

26. Kemp, A. Rearing of embryos and larvae of the Australian Lungfish, *Neoceratodus forsteri*, under laboratory conditions. *Copeia* **1981**, 776–784 (1981).

27. Nieuwkoop, P. D. & Faber, J. *Normal Table of Xenopus laevis (Daudin): A Systematical and Chronological Survey of the*

Development from the Fertilized Egg till the End of Metamorphosis 1 edn (Garland Publishing, 1994).

28. Hammerschmidt, M. et al. dino and mercedes, two genes regulating dorsal development in the zebrafish embryo. *Development* **123**, 95–102 (1996).

29. Kon, T. et al. The genetic basis of morphological diversity in domesticated goldfish. *Curr. Biol.* **30**, 2260–2274.e6 (2020).

30. Abe, G. et al. The origin of the bifurcated axial skeletal system in the twin-tail goldfish. *Nat. Commun.* **5**, 3360 (2014).

31. Watase, S. On the caudal and anal fins of goldfishes. *J. Sci. Coll. Imp. Univ. Tokyo* **1**, 247–267 (1887).

32. Yonei-Tamura, S. et al. Competent stripes for diverse positions of limbs/fins in gnathostome embryos. *Evol. Dev.* **10**, 737–745 (2008).

33. Smith, B. G. in *The Bashford Dean Memorial Volume: Archaic Fishes* Vol. 2 (ed. E. W. Gudger) Article VI, 333–506 (The American Museum of Natural History, 1937)

34. Cole, F. J. A monograph on the general morphology of the myxinoid fishes, based on a study of Myxine Part 1. The anatomy of the skeleton. *Trans. R. Soc. Edinb.* **41**, 749–788 (1905).

35. Larouche, O., Zelditch, M. L. & Cloutier, R. A critical appraisal of appendage disparity and homology in fishes. *Fish Fish.* **20**, 1138–1175 (2019).

36. Blom, H. New birkeniid anaspid from the Lower Devonian of Scotland and its phylogenetic implications. *Palaeontology* **55**, 641–652 (2012).

37. Kragl, M. et al. Cells keep a memory of their tissue origin during axolotl limb regeneration. *Nature* **460**, 60–65 (2009).

38. Richardson, L. R. *Neomyxine* n.g. (Cyclostomata) based on *Myxine biniplicata* Richardson and Jowett 1951, and further data on the

species. *Trans. R. Soc. N. Z.* **81**, 379–383 (1953).

39. Donoghue, P. C., Forey, P. L. & Aldridge, R. J. Conodont affinity and chordate phylogeny. *Biol. Rev. Camb. Philos. Soc.* **75**, 191–251 (2000).

40. Janvier, P. & Arsenault, M. The anatomy of *Euphanerops longaevus* Woodward, 1900, an anaspid-like jawless vertebrate from the Upper Devonian of Miguasha, Quebec, Canada. *Geodiversitas* **29**, 143–216 (2007).

41. Blom, H. A new anaspid fish from the Middle Silurian Cowie Harbour fish bed of Stonehaven, Scotland. *J. Vert. Paleontol.* **28**, 594–600 (2008).

42. Gai, Z. et al. Galeaspid anatomy and the origin of vertebrate paired appendages. *Nature* **609**, 959–963 (2022).

43. Ritchie, A. The late Silurian anaspid genus *Rhyncholepis* from Oesel, Estonia, and Ringerike, Norway. *Am. Mus. Novit.* **2699**, 1–18 (1980).

44. Mansfield, J. H. & Holland, N. D. Amphioxus tails: source and fate of larval fin rays and the metamorphic transition from an ectodermal to a predominantly mesodermal tail. *Acta Zool.* **96**, 117–125 (2015).

45. Tulenko, F. J. et al. Body wall development in lamprey and a new perspective on the origin of vertebrate paired fins. *Proc. Natl Acad. Sci. USA* **110**, 11899–11904 (2013).

46. Nuño de la Rosa, L., Müller, G. B. & Metscher, B. D. The lateral mesodermal divide: an epigenetic model of the origin of paired fins. *Evol. Dev.* **16**, 38–48 (2014).

47. Iwamatsu, T. Stages of normal development in the medaka *Oryzias latipes*. *Mech. Dev.* **121**, 605–618 (2004).

48. Kimmel, C. B., Ballard, W. W., Kimmel, S. R., Ullmann, B. & Schilling, T. F. Stages of embryonic development of the zebrafish. *Dev. Dyn.* **203**, 253–310 (1995).

49. Li, I. J., Lee, S. H., Abe, G. & Ota, K. G. Embryonic and postembryonic development of the ornamental twin-tail goldfish. *Dev. Dyn.* **248**, 251–283 (2019).

50. Tahara, Y. Normal stages of development in the lamprey, *Lampetra reissneri* (Dybowski). *Zool. Sci.* **5**, 109–118 (1988).

51. Ballard, W. W., Mellinger, J. & Lechenault, H. A series of normal stages for development of *Scyliorhinus canicula*, the lesser spotted dogfish (Chondrichthyes: Scyliorhinidae). *J. Exp. Zool.* **267**, 318–336 (1993).

52. Okuda, K. S. et al. lyve1 expression reveals novel lymphatic vessels and new mechanisms for lymphatic vessel development in zebrafish. *Development* **139**, 2381–2391 (2012).

53. Kwan, K. M. et al. The Tol2kit: a multisite gateway-based construction kit for Tol2 transposon transgenesis constructs. *Dev. Dyn.* **236**, 3088–3099 (2007).

54. Thermes, V. et al. I-SceI meganuclease mediates highly efficient transgenesis in fish. *Mech. Dev.* **118**, 91–98 (2002).

55. Thisse, C. & Thisse, B. High-resolution in situ hybridization to whole-mount zebrafish embryos. *Nat. Protoc.* **3**, 59–69 (2008).

56. Feitosa, N. M. et al. Hemicentin 2 and Fibulin 1 are required for epidermal-dermal junction formation and fin mesenchymal cell migration during zebrafish development. *Dev. Biol.* **369**, 235–248 (2012).

57. Ablooglu, A. J., Kang, J., Handin, R. I., Traver, D. & Shattil, S. J. The zebrafish vitronectin receptor: characterization of integrin alphaV and beta3 expression patterns in early vertebrate development. *Dev. Dyn.* **236**, 2268–2276 (2007).

58. Kaneko, H., Nakatani, Y., Fujimura, K. & Tanaka, M. Development of the lateral plate mesoderm in medaka *Oryzias latipes* and Nile tilapia

Oreochromis niloticus: insight into the diversification of pelvic fin position. *J. Anat.* **225**, 659–674 (2014).

59. Davis, M. C., Dahn, R. D. & Shubin, N. H. An autopodial-like pattern of Hox expression in the fins of a basal actinopterygian fish. *Nature* **447**, 473–476 (2007).

60. Lea, R., Bonev, B., Dubaissi, E., Vize, P. D. & Papalopulu, N. Multicolor fluorescent in situ mRNA hybridization (FISH) on whole mounts and sections. *Methods Mol. Biol.* **917**, 431–444 (2012).

61. Lea, R., Papalopulu, N., Amaya, E. & Dorey, K. Temporal and spatial expression of FGF ligands and receptors during *Xenopus* development. *Dev. Dyn.* **238**, 1467–1479 (2009).

62. Gilchrist, M. J. et al. Defining a large set of full-length clones from a *Xenopus tropicalis* EST project. *Dev. Biol.* **271**, 498–516 (2004).

63. Cerny, R. et al. Evidence for the prepattern/cooption model of vertebrate jaw evolution. *Proc. Natl Acad. Sci. USA* **107**, 17262–17267 (2010).

64. Kemp, H. A., Carmany-Rampey, A. & Moens, C. Generating chimeric zebrafish embryos by transplantation. *J. Vis. Exp.* **29**, 1394 (2009).

65. Metscher, B. D. MicroCT for comparative morphology: simple staining methods allow high-contrast 3D imaging of diverse non-mineralized animal tissues. *BMC Physiol.* **9**, 11 (2009).

Acknowledgements

We thank K. Poss and M. Tanaka for providing zebrafish and medaka *hand2* plasmids, respectively; Z. Gong for the transgenic fish line *TgBAC(hand2:EGFP)*; D. Traver for input on the first iteration of Dendra2 transgenics; Y.-K. Tea for input; and K. P. Lim of the Lee Kong Chian Natural History Museum, Singapore for assisting with the frilled shark specimen. We also thank the Nanyang Technological University Zebrafish facility, the NTU Optical Bio-Imaging Centre (NOBIC),

Colorado University (CU) Anschutz zebrafish care staff and all past and present members of our teams for continued support. This work was funded by the Industry Aligned Fund (IAF) Agency for Science, Technology and Research (grant to T.J.C. and K.-W.T.); Ministry of Education (MoE) Tier 3 (grant 2016-T3-1-005 to T.J.C., C.W. and H.M.); Ministry of Education (MoE) Tier 1 (grant 2016-T1-001-055 to T.J.C. and C.Z.); Ministry of Education (MoE) Tier 2 (grant MOE-T2EP30221-0008 to C.W.); the Company of Biologists (travelling fellowship to M.J.T.); the National Science Foundation (grants IOS-1853949 to M.C.D. and 2203311 to C.M.); the Swiss National Science Foundation Sinergia (grant CRSII5_180345 to C.M.); the Swiss Bridge Foundation (C.M.); Additional Ventures Single Ventricle Research Fund (SVRF) (grant 1048003 to C.M.); the University of Colorado School of Medicine Anschutz Medical Campus and the Children's Hospital Colorado Foundation (C.M.); the National Institutes of Health (NIH), National Institute of General Medical Sciences (grants 1T32GM141742-01 to H.R.M. and 3T32GM121742-02S1 to H.R.M.); Australian Research Council (discovery grant DP200103219 to F.J.T. and P.D.C); National Health and Medical Research Council (senior principal research fellow APP1136567 to P.D.C.); and the NIH (grant R35NS111564 to J.S. and M.E.B.).

Author information

Authors and Affiliations

1. Institute of Molecular and Cell Biology, A*STAR, Singapore, Singapore

   Keh-Weei Tzung, Monika J. Tomecka & Tom J. Carney

2. Lee Kong Chian School of Medicine, Nanyang Technological University, Singapore, Singapore

   Keh-Weei Tzung, Harsha Mahabaleshwar, Changqing Zhang & Tom J. Carney

3. Department of Pediatrics, Section of Developmental Biology, University of Colorado Anschutz Medical Campus, Aurora, CO, USA

Robert L. Lalonde, Karin D. Prummel, Hannah R. Moran & Christian Mosimann

4. Department of Molecular Life Sciences, University of Zurich, Zurich, Switzerland

Karin D. Prummel, Eline C. Brombacher & Christian Mosimann

5. Division of Biology and Biological Engineering, California Institute of Technology, Pasadena, CA, USA

Jan Stundl & Marianne E. Bronner

6. Faculty of Fisheries and Protection of Waters, University of South Bohemia in Ceske Budejovice, Vodnany, Czech Republic

Jan Stundl

7. Biology Department, Wesleyan University, Middletown, CT, USA

Amanda N. Cass

8. Department of Biological Sciences, Faculty of Science, National University of Singapore, Singapore, Singapore

Yao Le & Christoph Winkler

9. Division of Developmental Biology and Medicine, School of Medical Sciences, Faculty of Biology, Medicine and Health, University of Manchester, Manchester, UK

Robert Lea & Karel Dorey

10. CSIRO National Research Collections Australia, Australia National Fish Collection, Hobart, Tasmania, Australia

William T. White

11. School of Biosciences, University of Sheffield, Sheffield, UK

Henry H. Roehl

12. Australian Regenerative Medicine Institute, Monash University, Clayton, Victoria, Australia

Frank J. Tulenko & Peter D. Currie

13. EMBL Australia, Victorian Node, Monash University, Clayton, Victoria, Australia

Peter D. Currie

14. Division of Cell Matrix Biology and Regenerative Medicine, School of Biological Sciences, Faculty of Biology, Medicine and Health, University of Manchester, Manchester, UK

Enrique Amaya

15. Department of Physical and Biological Sciences, Western New England University, Springfield, MA, USA

Marcus C. Davis

Contributions

Experiments were conceived and designed by T.J.C., C.M., K.-W.T., M.E.B., M.C.D., E.A., P.D.C., C.W. and H.H.R. and were performed and analysed by K.-W.T., R.L.L., K.D.P., J.S., H.M., H.R.M., A.N.C., Y.L., R.L., K.D., M.J.T., C.Z., E.C.B., W.T.W. and F.J.T. The manuscript was written by K.-W.T. and T.J.C., with editing and input from C.M., J.S., A.N.C., M.C.D., F.J.T., P.D.C., R.L.L., K.D.P. and E.A.

Corresponding authors

Correspondence to Christian Mosimann or Tom J. Carney.

Ethics declarations

Competing interests

The authors declare no competing interests.

Peer review

Peer review information

Nature thanks William Bemis, Ann Burke, Brian Metscher and Oliver Stone for their contribution to the peer review of this work. Peer reviewer reports are available.

Additional information

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Extended data figures and tables

Extended Data Fig. 1 A non-paraxial mesoderm derived median fin fold.

a, Confocal image of 3 dpf *Tg(tbx16l:GAL4-VP16); Tg(UAS:Kaede)* embryo with paraxial mesoderm labelled by photoconverted Kaede (magenta). Unlabelled PAFF underscored by yellow bracket. **c**-**d**, Expression of Kaede in the dorsal caudal fin fold of 3 dpf *Tg(tbx16l:GAL4-VP16); Tg(UAS:Kaede)* is not due to localised de novo expression from the *tbx16l* promoter. Confocal images of 24 hpf trunks of both prior to (**b**) and after (**c**) UV photoconversion. Unconverted Kaede is in the green channel overlaid with converted Kaede in magenta. Region of paraxial mesoderm conversion and Nomarski image given in (**b**). At 3 dpf, converted cells can be seen in

the adjacent fin fold dorsally indicating Kaede is reporting lineage (**d**). **e**, Nomarski image of PAFF at 3 dpf showing presence of mesenchymal cells. **f**–**h**, Confocal images of the pre-anal fin fold (**f**), dorsal (**g**) and caudal (**h**) regions of the caudal fin fold of the *ET37* Enhancer Trap transgenic line at 3 dpf, indicating PAFF contains numerous mesenchymal cells which are morphologically comparable to mesenchyme of other larval fin folds. **i**–**j**, In situ hybridisation of fin mesenchyme markers *itgb3b* (**i**; 3 dpf) and *bmp1a* (**j**; 5 dpf) in both PAFFs (yellow) and caudal fin folds (cyan arrowhead). Scale Bars: 100 µm (**a**,**f**,**i**), 50 µm (**c**,**d**,**e**,**j**), 20 µm (**g**,**h**).

Extended Data Fig. 2 Expression of *hand2* in the pre-anal fin fold.

a–**b**, High (**a**) and low (**b**) magnification confocal images of *Tg(hand2:EGFP)* zebrafish embryos at 3 dpf (**a**) and 8 dpf (**b**) showing eGFP labelling of mesenchyme of the PAFF (yellow outline, **a**,**b**), but not the caudal fin fold (cyan outline, **b**). **c**–**d**, In situ hybridisation of *hand2* at 5 dpf showing expression of *hand2* only in the PAFF (**c**). Higher magnification with Nomarski optics indicates expression in mesenchyme (**d**). Scale Bars: 100 µm (**b**,**c**), 50 µm (**d**), 20 µm (**a**).

Extended Data Fig. 3 Loss of Hand2 leads to reduction of pre-anal fin folds but not caudal fin folds.

a–**l**, Lateral Nomarski images of uninjected WT (**a**,**c**,**e**,**g**), *han^{s6}* mutants (**b**,**h**) and *hand2* morphants (**d**,**f**) at 3 dpf (**a**–**d**) or 5 dpf (**e**–**h**). **i**–**l**, Confocal images of 3 dpf (**i**,**j**) and 5 dpf (**k**,**l**) *ET37* Enhancer Trap transgenic larvae, uninjected (**i**,**k**) or injected with a *hand2* morpholino (**j**,**l**). **m**,**n** Quantification of 3 dpf fin height of PAFFs (yellow) and caudal fin folds (cyan) in uninjected WT, *hand2* mutants (*han^{s6}*; **m**) and *hand2* morphants (**n**). Loss of Hand2 leads to significant reduction of PAFFs but not caudal fin folds. n = 8 (WT), n = 15 (*han^{s6}*), n = 5 (*hand2* MO). Location of measurements of PAFF and caudal fin shown in (**a**) and (**b**) by yellow and cyan lines respectively. ***: p = 0.000008; **: p = 0.008; ns: p = 0.12 (**m**); ns: p > 0.9999 (**n**). **o**, Quantification of mesenchymal cell numbers in pre-anal (yellow bars) and ventral caudal (cyan bars) fin folds at 3 dpf. There is a significant reduction of PAFF mesenchyme in *hand2* morphants compared

to WT, but ventral caudal fin mesenchymal cell numbers are unaffected. n = 5 (WT), n = 12 (*hand2* MO). ***: p = 0.000323; ns: p = 0.06. **m**–**o**: Data are presented as mean values with error bars representing standard deviations. In all cases n refers to biologically independent embryos. Two-sided Mann-Whitney test. Scale Bars: 100 µm (**d**,**f**,**j**,**l**), 200 µm (**b**,**h**).

Source data

Extended Data Fig. 4 Hand2 acts cell autonomously in PAFF mesenchyme.

a–**d**, Lateral confocal images of PAFF of 3 dpf embryos. Cells were transplanted into WT hosts from *ET37* embryos that were injected with *H2B-mCherry* mRNA (**a**,**b**) or co-injected with *H2B-mCherry* mRNA and 500 µM *hand2* MO (**c**,**d**). Fluorescent confocal images of mCherry (magenta) and eGFP (green) (**a**,**c**) are overlayed on Nomarski image (**b**,**d**). Arrowheads indicate transplanted cells from *hand2* MO injected donors (**c**). Scale Bar: 20 µm (**d**).

Extended Data Fig. 5 Lineage tracing of the LPM into paired fins and the PAFF.

a–**g**, Permanently labelled LPM cells can be seen within pectoral and pelvic fins and pre-anal fin fold (arrowheads). Confocal images of the pectoral fin at 4.5 dpf (**a**), adult pelvic fin (**b**–**d**) and pre-anal fin fold at 5 dpf (**e**–**g**) of *Tg(drl:creERT2); Tg(hsp70l:Switch)* transgenics following Hydroxytamoxifen treatment and heat-shock prior to imaging. Pelvic fins and PAFF are shown as transverse sections (**b**–**g**) and were fluorescently immunostained for eGFP (**b**,**e**), Zns5 (**c**) or Transgelin (**f**). Merged images are shown in (**d**,**g**). LPM-derived mesenchyme cells of the pectoral and pre-anal fin folds are indicated by white and yellow arrowheads respectively (**a**, **e**–**g**). LPM-derived Zns5-positive osteoblasts of the pelvic fins are indicated by white arrowheads (**b**–**d**). Scale Bars: 20 µm (**d**,**g**), 50 µm (**a**).

Extended Data Fig. 6 Pre-anal fin fold mesenchyme originates from posterior-most *drl*-positive cells.

a–**d**, Lateral (**a**) and ventral (**b**–**d**) views of 8ss (**a**–**b**) or 10ss (**c**–**d**) *drl:H2B-Dendra2* transgenic line either prior to (**a**–**b**) or following (**c**–**d**) photoconversion of posterior-most LPM (region outlined by dotted line) using UV laser illumination. **e**–**p**, Lateral confocal images of the PAFF region at 40 hpf (**e**–**g**, **k**–**m**) and 48 hpf (**h**–**j**, **n**–**p**) of a photoconverted embryo (**e**–**j**) and an unconverted control (**k**–**p**). Panels show Dendra2-red (**f**,**i**,**l**,**o**), Dendra2-green (**g**,**j**,**m**,**p**) and merged (**e**,**h**,**k**,**n**) channels. PAFFs are indicated by yellow arrowheads, and are visible in green channel in both converted and controls, but only converted embryos show red labelled PAFFs (**f**,**i**; n = 4 photoconverted fish). Magnified views of panels (**f**,**g**,**i**,**j**,**l**,**m**,**o**,**p**) are given in panels (**q**–**x**) respectively. Scale Bars: 200 µm (**b**), 50 µm (**d**,**x**), 100 µm (**k**).

Extended Data Fig. 7 LPM origin of PAFF mesenchyme is conserved in medaka, paddlefish and *Xenopus*.

a–**c**, In situ hybridisation of medaka *hand2* at Stage 39 (**a**) and Stage 40 (**b**–**c**) showing exclusive expression in pre-anal fin mesenchyme (yellow arrowhead, **b**), but not in the caudal fin mesenchyme (cyan outline). Ventral caudal fin imaged in (**a**). **d**–**e**, Lateral confocal images of PAFF of Stage 36 transgenic *Tg(-6.35drl:EGFP)* medaka embryo. Fluorescent confocal images of eGFP (**d**) are overlayed on Nomarski image (**e**). Faint eGFP perdurance is seen in nascent PAFF mesenchymal cells (yellow arrowheads and inset). **f**–**i**, Ventral (**f**,**h**) and lateral (**g**,**i**) confocal images of medaka embryo injected with DiI in the posterior LPM at Stage 20 (**f**,**h**) and traced to PAFF at Stage 40 (**g**,**i**). DiI within the mesenchyme indicated by yellow arrowheads, with higher magnification example shown in inset (**g**,**i**). DiI signal (**f**,**g**) is overlayed with Nomarski (**h**,**i**). Location of paraxial mesoderm (PM), Kupffer's vesicle (KV) and region of posterior LPM are indicated in green, red, and white respectively in (**h**). **j**–**l**, In situ hybridisation of *hand2* in Stage 38 (**j**) and Stage 39 (**k**–**l**) paddlefish embryos imaged laterally (**j**,**k**) or in transverse section (**l**). *hand2*-positive PAFF fin mesenchyme indicated by yellow arrowheads is located distally. Expression of *hand2* is also seen in the nascent pelvic fins at Stage 39 (green arrowheads **k**,**l**). **m**–**o**, Confocal images of a region of *Xenopus laevis* PAFF at NF Stage 42 injected with *-6.35drl:EGFP*. Lateral x-y view (**n**) is shown with orthogonal sections in the x-z plane (**m**) and y-z plane (**o**). Transient mosaic expression in the

interstitial mesenchyme is highlighted with yellow arrowheads (**m**–**n**). Scale Bars: 20 µm (**a**,**d**), 100 µm (**b**,**i**), 50 µm (**b** – inset, **l**,**o**), 200 µm (**h**,**j**,**k**).

Extended Data Fig. 8 Reduced Chordin leads to paired duplication of PAFF.

a–**b**, Low (**a**) and high (**b**) power Nomarski images of duplicated PAFFs (yellow arrowheads) in 5 dpf larvae injected with *chrd* morpholino. **c**–**j**, Confocal images of PAFFs from 8 dpf *Tg(hand2:EGFP)* either uninjected (**c**,**e**) or injected with *chrd* morpholino (**d**,**f**,**g**–**j**). Confocal projections (**c**,**d**,**i**,**j**) and total surface renderings (**e**,**f**,**g**,**h**) highlight duplicated PAFFs (blue and yellow, **f**) compared to single PAFF in uninjected larvae (grey, **e**). Imaging of left (**g**) and right (**h**) duplicated PAFFs of *chrd* morphants indicated eGFP positive mesenchymal cells populated both fin folds (yellow arrowheads, **i**,**j**). **k**– **l**, Lateral (**k**) and ventral (**l**) Nomarski images of the duplicated PAFFs of the Ranchu goldfish strain at 6 dpf. Duplicated PAFFs and ventral caudal fin folds indicated by yellow and cyan arrowheads, respectively. **m**–**n**, Lateral **(m)** and transverse (**n**) views of 7 dpf Ranchu larvae stained by in situ hybridisation for *hand2*. Expression in individual mesenchymal cells of the PAFF indicated by yellow arrowheads (**m**). Occurrence of three PAFFs with core *hand2* expression indicated by yellow arrowheads (**n**). Scale Bars: 100 µm (**a**,**e**,**g**), 50 µm (**b**), 500 µm (**k**), 200 µm (**l**), 20 µm (**m**,**n**).

Extended Data Fig. 9 Variable presence of a pre-anal fin in sharks.

a–**i**, The Epaulette shark does not possess a PAFF. Whole mount (**a**,**d**,**g**) and virtual cross sections from microCT scans (**b**–**c**, **e**–**f**, **h**–**i**) of the Epaulette shark *Hemiscyllium ocellatum*, at the developmental stages of 25 (**a**–**c**), 27 (**d**–**f**), and 30 (**g**–**i**). Virtual cross sections for each stage show the region either anterior (Pre-anal; **b**,**e**,**h**), or posterior (Post-anal; **c**,**f**,**i**) to the developing cloaca (white arrow, **a**,**d**,**g**). The dashed lines (**a**,**d**,**g**) indicate approximate pre-anal (magenta) and post-anal (cyan) microCT section locations (separate specimens). The developing pelvic and median fin folds are indicated by green and cyan arrowheads respectively (**b**–**c**, **e**–**f**, **h**–**i**). **j**–**m**: The tropeic folds found in a pre-anal position of the adult frilled shark,

Chlamydoselachus anguineus, shown in lateral (**j**) and ventral (**k**) views (yellow arrowheads) (adult male specimen, ZRC 54430 - LKC Natural History Museum, Singapore). **l**–**m**, Ventral view of another adult frilled shark exhibits partially paired tropeic folds (yellow arrowheads) in anterior half (**l**), which then merges to become singular (yellow arrowhead) in posterior half (**m**) (anterior: left, CSIRO H 7115-01, adult male 1310 mm TL, Tasmania, Australia). Scale Bars: 1 mm (a,d,g), 5 cm (**k**).

Supplementary information

Reporting Summary

Peer Review File

Supplementary Video 1

A 3D reconstruction of the duplicated PAFFs in *chrd* morphants. 3D rotation, surface rendering and fly-through of 8 dpf *Tg(hand2:EGFP)* larvae uninjected (left larva) or injected with *chrd* morpholino (right larva), with eGFP-positive mesenchyme visible in all PAFFs.

Source data

Source Data Extended Data Fig. 3

Rights and permissions

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons

licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by/4.0/.

Reprints and Permissions

About this article

Cite this article

Tzung, KW., Lalonde, R.L., Prummel, K.D. *et al.* A median fin derived from the lateral plate mesoderm and the origin of paired fins. *Nature* **618**, 543–549 (2023). https://doi.org/10.1038/s41586-023-06100-w

- Received: 05 August 2022
- Accepted: 19 April 2023
- Published: 24 May 2023
- Issue Date: 15 June 2023
- DOI: https://doi.org/10.1038/s41586-023-06100-w

Share this article

Anyone you share the following link with will be able to read this content:

Get shareable link

Sorry, a shareable link is not currently available for this article.

Copy to clipboard

Provided by the Springer Nature SharedIt content-sharing initiative

This article was downloaded by **calibre** from https://www.nature.com/articles/s41586-023-06100-w

- Article
- Open Access
- Published: 07 June 2023

Northwest African Neolithic initiated by migrants from Iberia and Levant

- Luciana G. Simões,
- Torsten Günther,
- Rafael M. Martínez-Sánchez,
- Juan Carlos Vera-Rodríguez,
- Eneko Iriarte,
- Ricardo Rodríguez-Varela,
- Youssef Bokbot,
- Cristina Valdiosera &
- …
- Mattias Jakobsson

Show authors

Nature **volume 618**, pages 550–556 (2023)

- 16k Accesses
- 1 Citations
- 403 Altmetric
- Metrics details

Abstract

In northwestern Africa, lifestyle transitioned from foraging to food production around 7,400 years ago but what sparked that change remains unclear. Archaeological data support conflicting views: (1) that migrant European Neolithic farmers brought the new way of life to North Africa[1,2,3] or (2) that local hunter-gatherers adopted technological innovations[4,5]. The latter view is also supported by archaeogenetic data[6]. Here we fill key chronological and archaeogenetic gaps for the Maghreb, from Epipalaeolithic to Middle Neolithic, by sequencing the genomes of nine individuals (to between 45.8- and 0.2-fold genome coverage). Notably, we trace 8,000 years of population continuity and isolation from the Upper Palaeolithic, via the Epipaleolithic, to some Maghrebi Neolithic farming groups. However, remains from the earliest Neolithic contexts showed mostly European Neolithic ancestry. We suggest that farming was introduced by European migrants and was then rapidly adopted by local groups. During the Middle Neolithic a new ancestry from the Levant appears in the Maghreb, coinciding with the arrival of pastoralism in the region, and all three ancestries blend together during the Late Neolithic. Our results show ancestry shifts in the Neolithization of northwestern Africa that probably mirrored a heterogeneous economic and cultural landscape, in a more multifaceted process than observed in other regions.

Main

North Africa's geographic location, centred between the vast Saharan desert, the fertile Near East and Mediterranean Europe, has resulted in a complex human history in the area[7,8]. The fossil record suggests long-term hominid and human presence[9], although continuity over the past 100,000 years cannot be deduced due to the fragmented nature of the record. In the Late Pleistocene, 15,000 years ago, the remains of foragers excavated in Morocco show a distinct genetic make-up intermediate between contemporary Levantine foragers and sub-Saharan African populations[10]. Current-day North Africans are largely related to Eurasian populations, which was probably caused by 'back-to-Africa' migrations[7].

Both archaeological records and archaeogenomic data show that Neolithic farmers (genetically distinct from European foragers) dispersed from the

northern Levant and Anatolia to the Mediterranean islands, Italian peninsula and Iberia[11,12,13,14,15,16,17,18]. Mediterranean coastal routes have long been recognized in the archaeological record as an important part of the Neolithic expansion in Europe. In the western Mediterranean, Impressed Ware technology—and further the Cardial Horizon—spread along the European mainland coast and islands to reach the Iberian peninsula, where both phenomena are present at 7,550 calibrated years before the present (cal bp) (refs. [19,20]).

Whereas some studies support a simultaneous appearance of the Neolithic in northwestern Africa (Eastern Rif, Ifri Oudadane site) and Iberia around 7,550 cal bp (ref. [21]), the earliest evidence for pottery, domestic cereals and husbandry is found in northern Morocco approximately two centuries later at Kaf Taht el-Ghar (KTG) around 7,350 cal bp (refs. [2,3,22,23]). Although Early Neolithic material culture and the first domestic mammals and pulses suggest a connection to Iberia[1,2,3], the extent and legacy of these connections remain unclear. However, the first genomic analysis of Early Neolithic farmers from northwestern Africa (from the site Ifri n'Amr o'Moussa (IAM) in central Morocco) shows no traces of admixture with European Neolithic farmers. Instead, it shows long-term population continuity since the Upper Palaeolithic in the region[6]. This result aligns with the hypothesis that the Neolithic transition in northwestern Africa was initiated by local Epipalaeolithic communities adopting technological innovations[4,5], such as those found at IAM: impressed Cardial-like ceramics, similar to those present throughout the western Mediterranean Neolithic Europe, and domestic cereals (for example, a grain of *Hordeum vulgare* dated around 7,050 cal bp)[2]. This pattern implies a Neolithization process that contrasts markedly with that of Europe, where it has been established that agriculture was introduced by the west- and northward demic diffusion of Anatolian early farmers[11,12]. The local development, or acculturation, of the North African Neolithic is further supported by signs of increasingly sedentary Epipalaeolithic groups developing strategies for resource management, such as the exploitation of wild plants and pottery[1,4,24,25,26]. Rapid climatic changes favoured mobile herding[27] and, whereas it has been hypothesized that cattle were independently domesticated in the Sahara[28], radiocarbon data suggest a gradual introduction of pastoralism in the Sahara in a

southwestwards direction 7,000–6,000 cal bp, possibly from the Near East[29,30].

Whereas palaeogenomic studies on the European Mediterranean Neolithic transition are abundant[15,31,32,33], North Africa has been the focus of only a single study that generated human genetic data from one Early and one Late Neolithic site[6], leaving substantial gaps in the chronology of events. It is evident that the site of IAM shows a Neolithic lifestyle and an absence of European Neolithic ancestry, but whether this was an independent development or the inspiration came from other groups in northwestern Africa or across the Mediterranean Sea remains unclear. Hence, the timeline and processes involved in the Neolithization of the region, the nature and dynamics of different economies in North Africa and the role they may have played in the broader European Neolithic remain understudied and controversial.

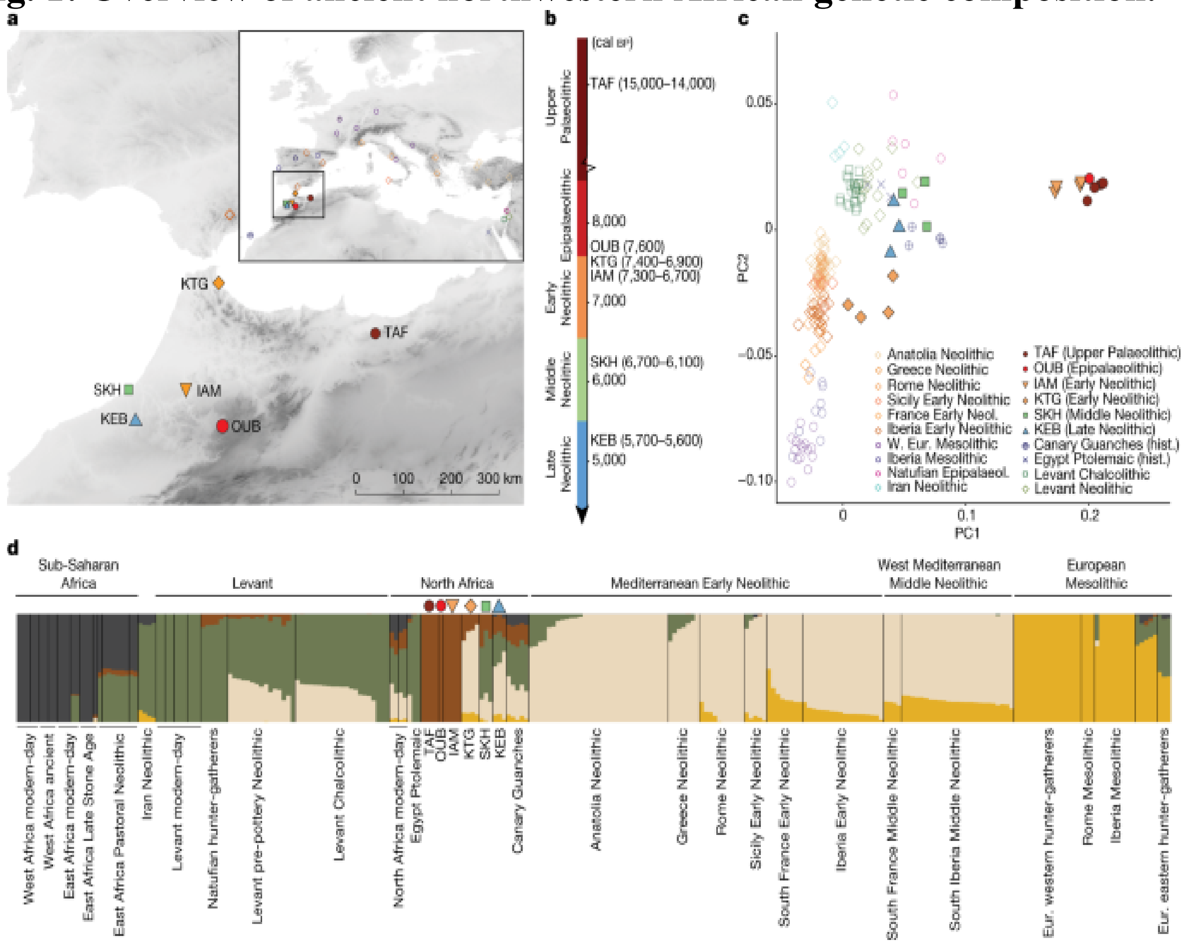
In this study we investigate a time series of human remains from four archaeological sites spanning the Epipalaeolithic to Middle Neolithic in current-day Morocco: the Epipalaeolithic site of Ifri Ouberrid (OUB), the Early Neolithic sites of IAM and KTG and the Middle Neolithic cemetery of Skhirat-Rouazi (SKH), co-analysed with previously published genetic data from that region[6,10]. By sequencing the genomes of nine individuals excavated from these four archaeological sites, we can demonstrate that the Neolithic transition in northwestern Africa was ignited by migration of Neolithic farmers from Mediterranean Europe.

We generated genomic sequence data from nine ancient individuals from modern-day Morocco (Table 1), ranging in genome coverage from 45.75- to 0.017-fold, including five individuals with more than onefold coverage and three with more than ninefold. Chronologically the data span more than 1,000 years, covering the Late Epipaleolithic ($n = 1$), Early Neolithic ($n = 5$) and Middle Neolithic ($n = 3$). Two Early Neolithic sites were studied—KTG ($n = 4$) and IAM—where we co-analysed the newly generated genomic data of one individual and those previously reported[6] (Fig. 1a,b). DNA libraries were generated from DNA extracts obtained from bones and teeth and subsequently shotgun sequenced on an Illumina platform. All libraries presented the degradation patterns expected from ancient DNA, including

short fragment sizes and cytosine deamination at read ends (Supplementary Fig. 1). Contamination estimates were generally low for both the nuclear genome and mitochondria except for individual skh003, which showed 10–16% nuclear contamination (Table 1). To assess the relationship of the ancient northwestern African individuals to other ancient and present-day West Eurasian and African populations, we co-analysed our data with relevant ancient (Supplementary Data 2) and current-day groups from Africa, the Middle East and Europe[34].

Table 1 Summary information of archaeological and newly generated genomic data from the ancient individuals reported in this study

Fig. 1: Overview of ancient northwestern African genetic composition.



a, Geographic location of investigated archaeological sites. Symbol legend given in **c**. The map was generated using the open source QGIS Geographic Information System, http://qgis.osgeo.org. **b**, Chronological representation of the investigated archaeological time periods of northwestern Africa, with

each site's radiocarbon-dated timeline indicated. **c**, Enlarged view of a PCA plot (Supplementary Fig. 3) with focus on the ancient individuals analysed. Each projected ancient individual is represented by a coloured symbol. W. Eur., West European; hist., historical. **d**, Estimated ancestry proportions for relevant African, Middle Eastern and European (Eur.) modern-day and ancient individuals (assuming five ancestry components; additional results are presented in Supplementary Fig. 4). Pre-Neolithic and Neolithic northwestern African populations/individuals are highlighted by the same symbols used in **a** and **c**.

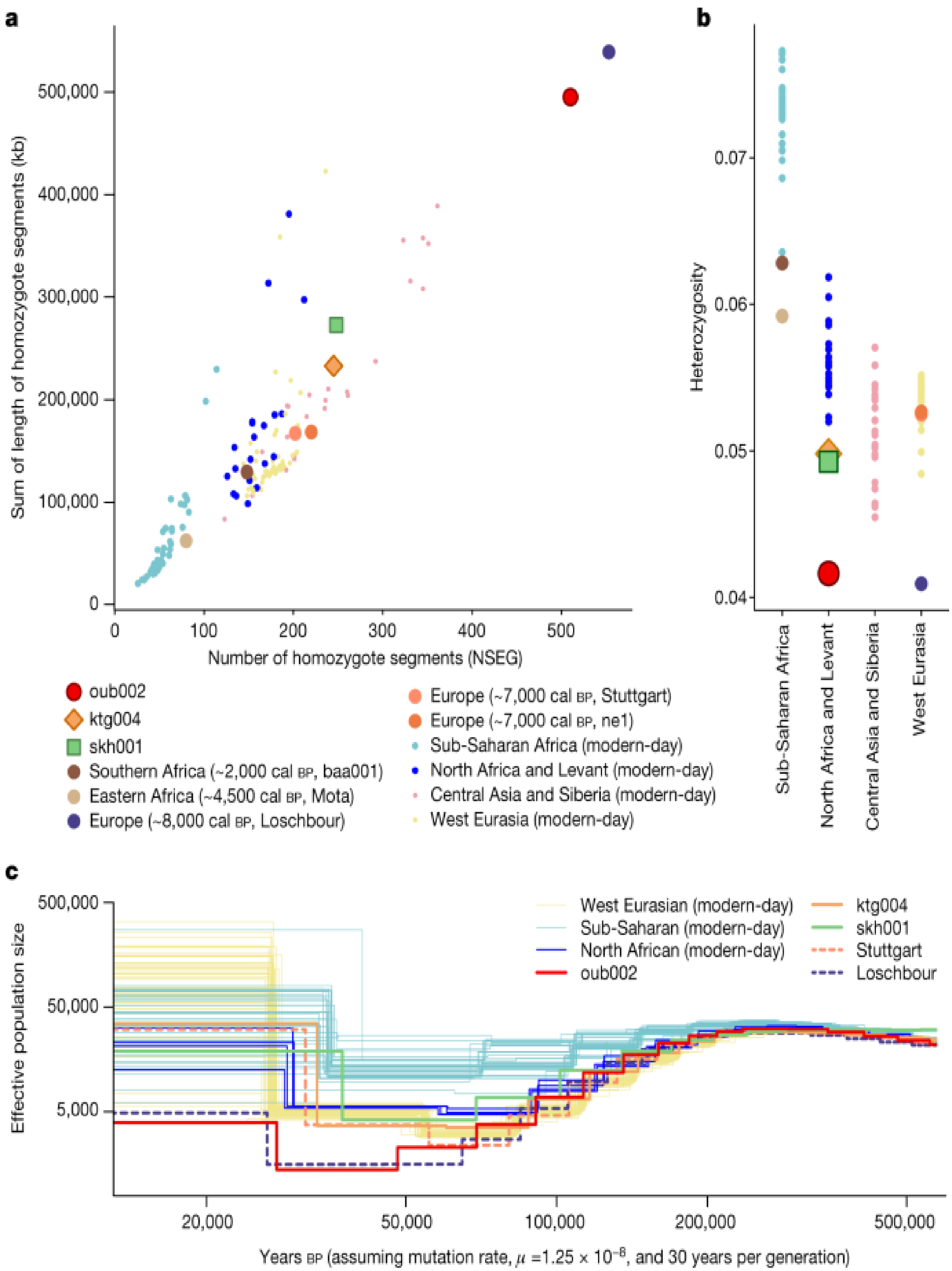
Eight thousand years of population continuity

From the Upper Palaeolithic people of Taforalt (TAF) via the Epipalaeolithic at OUB to the Early Neolithic at IAM, we observe the persistence of the unique genetic make-up that existed in northwestern African inhabitants 15,000 years ago (Fig. 1c,d and Supplementary Fig. 5), and possibly even further back in time. The Epipalaeolithic individual oub002, dating to 7,660–7,506 cal bp, is genetically very similar to individuals from TAF (15,086–14,046 cal bp)[35] and Early Neolithic individuals from IAM (7,316–6,679 cal bp; Fig. 1)[6,36]. The genome of Oub002 demonstrates a marked population continuity in northwest Africa with no substantial gene flow across the Mediterranean Sea for at least 7,000 years across the Epipalaeolithic (Fig. 1c,d), linking the Maghrebi genetic ancestry found in the Upper Palaeolithic to the Early Neolithic individuals at IAM.

The Maghrebi lineage shows outstandingly low genetic diversity[6,10] (Fig. 2b and Supplementary Fig. 9) and long and frequent runs of homozygosity (RoH) (Fig. 2a), probably as a consequence of long-lasting isolation. By investigation of the 45.8-fold genome of oub002 we show that ancient northwestern Africans went through a severe population bottleneck. Until some 70,000–60,000 years ago the effective population size (N_e) changes of oub002 follow a pattern similar to that of Eurasian populations with a relatively small effective population size reached 50,000 years ago (Fig. 2c), which is consistent with the Maghrebi lineage being related to the populations that migrated out of Africa. Interestingly, modern-day Eurasians and North Africans, as well as Neolithic Eurasians effective population size remained at around 5,000 until about 30,000 years ago but the effective

population size of the Maghrebi lineage continues to decrease and reached its lowest point ($N_e \approx 1{,}400$) between 50,000 and 27,000 years ago during the peak of the Last Glaciation. Remarkably similar patterns are observed for the Mesolithic western hunter-gatherers (WHG) of Europe (represented by Loschbour in Fig. 2c), for which low diversity measures have been attributed to high levels of background relatedness and autozygosity due to small population size[37].

Fig. 2: Measures of genetic diversity in ancient (northwestern and sub-Saharan) Africans and Eurasians, computed using diploid calls from higher-coverage (over ninefold genome coverage) individuals.



a–**c**, Ancient individuals (including oub002, ktg004 and skh001) are compared with modern-day individuals from geographically corresponding

regions. **a**, Runs of homozygosity. ne1, Neolithic European 1; NSEG, number of homozygote segments. **b**, Heterozygosity, calculated from the number of variable positions per individual divided by the number of single-nucleotide polymorphism (SNP) sites per individual. **c**, Effective population size over time, as inferred by pairwise sequentially Markovian coalescent, for three ancient northwestern Africans with over ninefold genome coverage, as well as for a Mesolithic European individual (Loschbour) and a Neolithic European individual (Stuttgart), and modern-day individuals for comparison.

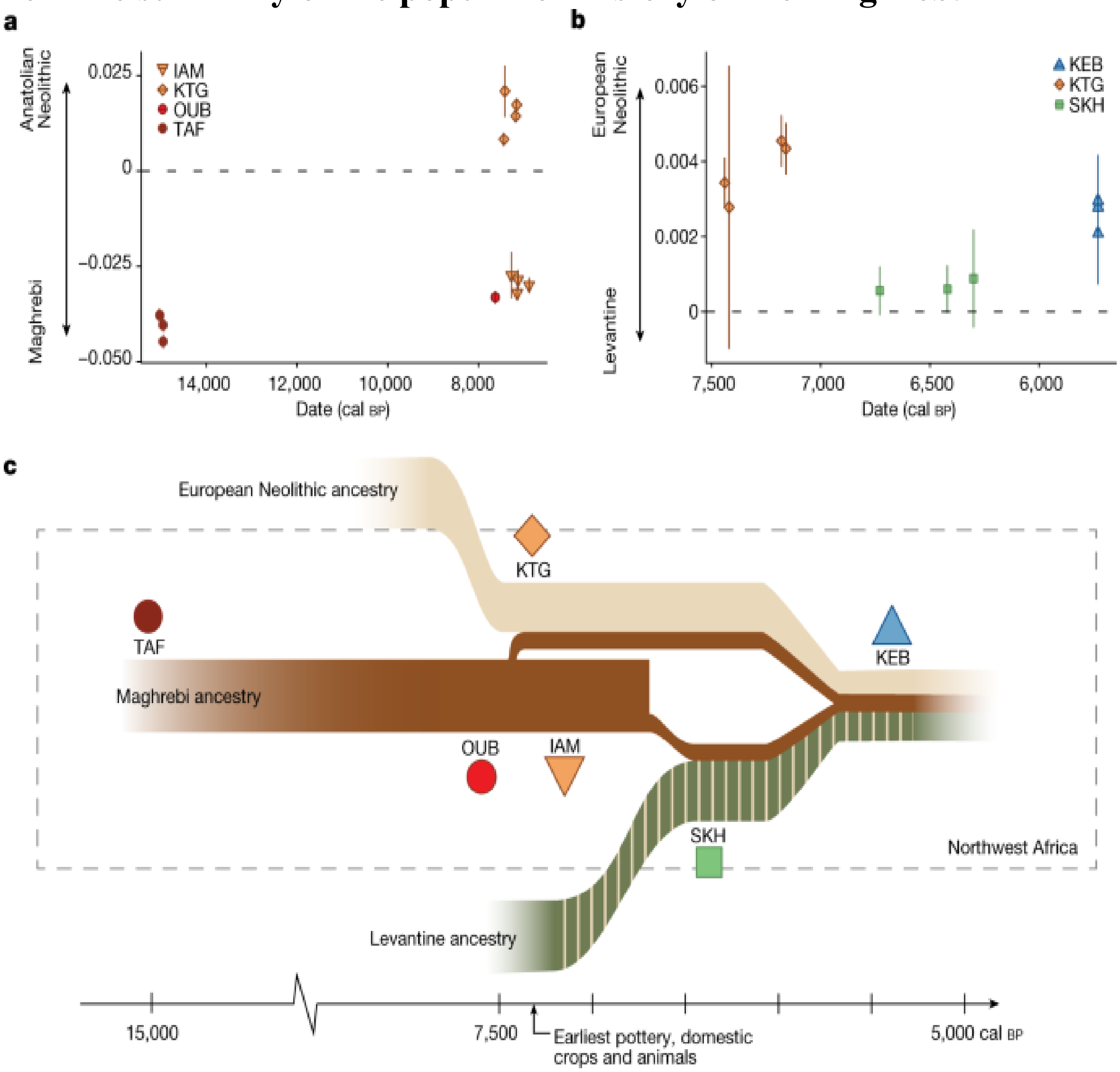
European farmers induce Neolithization

At the site of IAM, a multitude of artefacts representing the Neolithic package have been identified. However, it has been shown that the people living at IAM show autochthonous Maghrebi ancestry[6] and were the descendants of earlier (Upper Palaeolithic and Epipaleolithic) northwestern African groups (Fig. 1c,d). These two observations support the view that the first stage of the Neolithic transition in Morocco was driven by local populations adopting technological innovations based on contacts across the Mediterranean[2].

The Early Neolithic site of KTG, located on the North African Mediterranean coast near the Gibraltar strait (Fig. 1a), predates and partly overlaps in time with IAM[2] (Table 1). At KTG a full Neolithic assemblage is found, including a diversity of cultivated cereals, domestic mammals and cardial ceramics[38,39]. In contrast to the people at IAM, those at KTG are genetically similar to European Early Neolithic populations (Figs. 1c,d and 3a). Interestingly, all four KTG individuals show admixture (15.4–27.4%) with local North African groups (Fig. 1d), consistent with significantly positive values for the f_4 test of admixture (KTG, Mediterranean EN; TAF, Mbuti) (Supplementary Data 7). Furthermore we identify a small proportion of WHG ancestry in KTG (Fig. 1d), consistent with the observation of Early Neolithic Europeans carrying WHG ancestry[14,15,31,33,40]. A population history model for the KTG people with 72 ± 4.4% Anatolian Neolithic ancestry, 10 ± 2.6% WHG ancestry and 18 ± 3.3% Maghrebi ancestry is consistent with the data (qpAdm, P = 0.193). Taken together, these results

suggest a European Neolithic origin of KTG farmers whose ancestors dispersed from Anatolia throughout Europe, admixing with European hunter-gatherers on their path to southwestern Europe[33,40] before crossing the Mediterranean to North Africa. The presence of European hunter-gatherer ancestry excludes the possibility that Early Neolithic migrants exclusively followed North African Mediterranean shores from Anatolia or the Levant.

Fig. 3: Genetic affinities of Stone Age northwestern Africans and schematic summary of the population history of the Maghreb.



a, Genetic affinity of analysed Stone Age northwestern African individuals, polarized between Neolithic Anatolia and Maghrebi ancestry, using the f_4

test of the form f_4(Anatolia Neolithic, TAF011; Stone Age northwestern African individuals, Mbuti). **b**, Genetic affinity of Stone Age northwestern African individuals, polarized between Early Neolithic Iberian and Levantine ancestry using f_4 test of the form f_4(Iberia Early Neolithic, Levant Chalcolithic; Neolithic northwestern African individuals, Mbuti). **a**,**b**, Each symbol represents a single individual f_4 value. Error bars indicate ±2 s.e., computed with a block jack-knife approach (5 Mb blocks weighted by the number of SNPs). **c**, Summary of inferred population history of the Stone Age Maghreb.

Iberian Early Neolithic (both as a whole and regionally) was found to be the best source population for the European ancestry in KTG, followed by Sicily Stentinello Early Neolithic (Supplementary Data 9). This is consistent with low levels of genetic differentiation in Cardial Ware-associated groups along the European shores of the Mediterranean Sea[41], confirmed by direct radiocarbon dates showing that Impressed Ware farmers expanded rapidly across the western Mediterranean[3,19,42].

It has been debated whether European farmers crossed from Iberia to Morocco[2,3] or whether earlier crossings of the Mediterranean would have happened, through the Sicilian–Tunisian Strait followed by a Maghrebi route of expansion[4,43]. Direct comparisons of Early Neolithic farmers from Sicily and Iberia as ancestors of KTG farmers provide stronger evidence for an Iberian Neolithic origin (Supplementary Data 9 and Supplementary Information 8), but we cannot exclude some contribution from Sicilian farmers. Genetic data are consistent with the most parsimonious explanation for archaeological evidence on the Neolithic transition in northwestern Africa: the crossing from southern Iberia by Iberian Neolithic farmers[2,23]. The close geographical proximity between southern Iberia and the Tangitana Peninsula adds strength to this observation whereas the lack of reliable archaeological evidence of early domestic elements in relevant sites along the eastern Maghreb and Tunis, including sites with pottery and obsidian from Pantelleria Island, undermines the Sicily–Tunis crossing hypothesis[3]. Interestingly, gene flow from North Africa was found only in Mediterranean European individuals much later, from around 4,500 years ago[31,44].

Different individuals from KTG date to slightly different time periods. We find a twofold larger proportion of Maghrebi ancestry in earlier KTG individuals (roughly 25%, ktg001 and ktg005, approximately 7,429–7,267 cal bp) than in later ones (about 13%, ktg004 and ktg006, around 7,247–6,945 cal bp) (Fig. 1d). This coincides with an increase in European Neolithic ancestry, shown by the significantly negative result for f_4(KTG earlier, KTG later, Iberia Early Neolithic, Mbuti; z-score = −5.01). Approximately one quarter of Maghrebi ancestry in early KTG suggests that they represent at least the second generation of interbreeding between the groups. We estimated the time of admixture using two approaches based on ancestry covariance patterns and linkage disequilibrium decay, using Iberia or Sicily Early Neolithic and TAF as admixture sources. Both methods date the contact within the last six to 13 generations (Supplementary Information 8), suggesting that mixing between groups occurred for a few hundred years, which is consistent with analysis of pottery style that points to the first contact at 7,500–7,400 cal bp (ref. [23]).

Kaf Taht el-Ghar farmers had slightly lower genetic diversity levels and greater RoH than most Early Neolithic European populations (Fig. 2a,b and Supplementary Fig. 9). The Maghrebi ancestry carried by KTG people shows markedly lower diversity and more extensive RoH, and is probably the cause of the reduction in overall diversity. Archaeological evidence suggests that Early Neolithic farming was restricted to enclaves in westernmost Maghreb, possibly due to climatic constraints to the south[4,22]. This could have limited the potential of these groups to recover from an initial founder effect.

Overall, the genetic patterns of local interaction between different groups in northwestern Africa are comparable to those found in Europe: farmers assimilated local foragers’ ancestry in a unidirectional admixture process. Cases of hunter-gatherer communities adopting certain elements of the Neolithic have been described in Europe[11,14,45]. However, the northwestern Africa Neolithization process involved the notable survival of genetically unadmixed local populations (represented by IAM), despite coexisting for at least 300 years with foreign farming communities (KTG), and still adopted several elements of the Neolithic ways of living from them. Whereas the archaeological findings in IAM and KTG point to the exchange of ideas

between groups and support an acculturation process of foraging communities[1,4], our genetic data show that the exchange of genes was unidirectional.

Influx of Levantine ancestry

Another, distinct, ancestry was introduced to northwestern Africa during the Middle Neolithic. All individuals from SKH show large proportions of a genetic component maximized in individuals from Neolithic and Chalcolithic Levant, Ptolemaic Egypt and modern-day Near Eastern populations (Fig. 1d). The ancestry in SKH can be modelled as a two-way admixture between Levant Neolithic populations (roughly 76.4 ± 4.0%) and local northwestern Africans (represented by TAF; 23.6 ± 4.0%). If a European Neolithic (for example, from Iberia) additional source population is added, the model is rejected.

Because this Neolithic Levantine ancestry has not been observed on the European side of the Mediterranean during the Neolithic, it probably represents an independent expansion of people from the Levant into North Africa. Migrations from the Levant to eastern Africa have been identified for Neolithic pastoralist individuals around 4,000 years ago, who are presumed descendants of unsampled northeastern African populations associated with the spread of Saharan pastoralism[46]. Both in SKH and eastern African Neolithic pastoralists, Levantine ancestry is admixed with local ancestries (Fig. 1d, Supplementary Information 8 and Supplementary Data 12). The arrival of this Levantine ancestry coincides with the appearance of a new ceramic tradition in northern Morocco, often characterized by cord-impressed motifs (‘roulette’ or wavy line), like the grave goods at Skhirat belonging to Ashakar Ware pottery[47,48]. In parallel, cattle pastoralism was expanding in the current Sahara territory[30,47] and Afro-Asiatic language groups spread throughout the whole of North Africa[22].

Our analyses show that the Levantine-associated component also remains in the Maghreb during the Late Neolithic in individuals from Kehf el Baroud (KEB) and in the Guanches of the Canary Islands (around 1,000 cal bp; Fig. 1c,d)[6,49]. Individuals from these sites are shifted towards ancient Levantine populations on the principal component analysis (PCA) space (Fig. 1c). This

highlights the complex demographic processes that took place in northwestern Africa, in contrast to the gradual increase in hunter-gatherer ancestry described in Middle and Late Neolithic Europe[32,33,40].

The Late Neolithic individuals from KEB can be modelled as a mix of ancestries already present in northwestern Africa during the Early Neolithic and Middle Neolithic, suggesting that there were no waves of substantial migration into this region between the Middle Neolithic and Late Neolithic (Supplementary Information 8 and Supplementary Data 13).

Conclusion

The complex population structure in modern-day northwestern Africa has been linked to various historical events, such as the Arab expansion[7,8]. However, our detailed chronology and high-resolution genomic data provide a new understanding of these prehistoric processes in the Maghreb and unveil a rich and diversified genetic substrate with Neolithic origin. First, human populations in northwestern Africa show genetic continuity and isolation since the Upper Palaeolithic, from at least 15,000 to around 7,500 years ago, when this period of isolation was interrupted by the migration of European Early Neolithic groups introducing farming practices. Hence, despite a relatively small geographic distance between southern Iberia and northwestern Africa (the distance today is only 13 km across the Gibraltar straight), and the fact that both regions were populated by foragers for many millennia prior to the Neolithic, gene flow across the Mediterranean Sea was not established until the Early Neolithic. The newcomers brought new ways of life, farming practices, domestication and pottery traditions that were subsequently adopted by local populations. Our results show that the Neolithization process in northwestern Africa was ignited by migrant Neolithic Europeans, but that local groups (at least the individuals analysed at IAM) adopted some of these practices without mixing with the newcomers. Two genetically distinct groups coexisted in close proximity in the region. Interestingly, cultural and technological knowledge appear to have been transferred mainly from European Neolithic farmers to local groups (for example, at IAM) whereas genetic ancestry flowed only from local groups to the incoming farmers, such as the population of KTG. Furthermore, in the Middle Neolithic a new ancestry with an eastern origin

is detected in northwestern Africa. This ancestry indicates new migrating groups, potentially associated with Sahara pastoralists, which admixed with local groups (Fig. 3c).

The various waves of migration and admixture into northwestern Africa during the Neolithic possibly resulted in a heterogeneous economic and cultural landscape in that region—a mosaic of groups that included incoming farmers from Iberia, foragers adopting farming practices and eastern pastoralists admixing with local people. Most of these groups showed reduced effective population size and lower diversity than the contemporary populations in Europe (Fig. 2), suggesting that population sizes remained modest throughout the Neolithic. These patterns were probably caused by periods of isolation, which may have contributed to the distinct genetic ancestry seen in the Maghreb today. A recent study from the Iron Age suggests that northwestern Africa remained home to a diverse set of groups throughout prehistory[50], making this part of the world one of the most unique places to have been studied with the archaeogenomic toolkit.

Methods

Detailed descriptions for each section can be found in Supplementary Information.

Archaeological sampling

The ancient human remains analysed in this study derive from a scientific cooperation agreement between INSAP, La Trobe and Uppsala Universities. Complete bone and teeth elements were brought to the ancient DNA facility in Uppsala, Sweden for further cleaning and sampling.

Radiocarbon dating

All individuals investigated were directly radiocarbon dated at the Tandem Laboratory, Uppsala, except for ktg001, who was dated at the Beta Analytic Carbon dating laboratory, and iam004, who’s date was obtained from ref. [6]. Radiocarbon calibration for newly reported and relevant previously published dates was performed using Oxcal v.4.4 and the IntCal20 dataset[41].

Ancient DNA retrieval

Human remains were sampled in dedicated clean-room laboratories at Uppsala University, Sweden after a series of stringent procedures aimed at minimization of bone and tooth surface contamination. Thirty to sixty milligrams of bone powder or solid pieces of bone material were used for DNA extraction either following ref. [51], with adaptations as described in ref. [15], or following ref. [52], with adaptations to the binding buffer, and an initial predigestion step with 1 ml of 0.5 M EDTA pH 8.0 for 30 min at 37 °C[53]. Sample digestion was performed overnight with 1 ml of 0.45 M EDTA pH 8.0 and 0.2 mg ml^{-1} proteinase K. Double-stranded, blunt-end-repaired DNA libraries were built with ligated P5 and P7 adaptors[54]. After assessment of DNA authenticity, quality and quantity (by estimation of endogenous DNA content, post mortem deamination patterns and fragment size distribution), the remaining DNA extract (for samples with over 1% proportion of human DNA) was used to build four to six additional double-stranded DNA libraries; for extracts with roughly 5% endogenous human content or more, 15–20 µl of DNA extract was treated with uracil DNA glycosylase (UDG) for double-stranded library building[55]. Libraries were PCR amplified using a unique 7 bp indexed primer[54,56] in either four reactions of 25 µl or two of 50 µl, with the application of 12–20 PCR cycles depending on previous qPCR quantification cycle indication. Two extraction negative controls, two library negative controls and one PCR negative control were included per sample batch. PCR reactions were pooled and purified with AMPure XP beads (Agencourt, Beckman Coulter). Library quality was checked by electrophoresis on Tapestation (Agilent High Sensitivity D1000 ScreenTape, Agilent) and DNA concentration was quantified using a Qubit dsDNA HS (High Sensitivity) Assay Kit (Invitrogen). Equimolar pools of amplified and purified libraries were sequenced on Illumina HiSeq X at the SNP & SEQ Technology Platform in Uppsala. To reach higher coverage, between four and ten libraries were pooled equimolarly and sequenced to depletion.

Bioinformatics data processing and authentication

Data were demultiplexed according to the indexed primer sequence and adaptors were trimmed with either MergeReadsFastQ_cc.py[57] or Adapter

Removal v.2.1.7 (ref. [58]). Forward and reverse paired-end reads were merged when an overlap of at least 11 bp was found. Mapping against the human reference genome build 37 (hs37d5) was done using Burrows–Wheeler aligner 0.7.13 (ref. [59]). For each library we merged bam files resulting from all resequencing rounds using SAMtools merge v.1.5 (ref. [60]). We then separately merged data from UDG-treated and untreated libraries for each individual and used data from the former for subsequent analysis, except for individuals ktg001 (for which only non-UDG data were generated) and iam004 (for which both treated and untreated data were merged and processed as non-UDG treated for downstream analysis). We used a modified version of FilterUniqSAMCons_cc.py[57] to ensure random choice of bases to collapse reads with identical start and end positions into a consensus, thereby removing PCR duplicates. Reads shorter than 35 bp and more than 10% mismatches to the human reference genome were filtered out.

Contamination, sex determination, uniparental markers and kinship analyses

Sample contamination estimates were obtained using three different methods based on the mitochondrial genome[61], on the X chromosome in males[62] and on nuclear data[63] (Supplementary Data 3). The ratio of coverage of X and Y chromosomes relative to autosomes was used to determine the biological sex of each individual[64]. We generated mitochondrial consensus sequences using SAMtools 1.5 mpileup and vcfutils.pl[60,65]. Base (BQ) and mapping quality (MAPQ) scores were set to MAPQ > 30 and BQ > 30, and only sites with at least threefold coverage were used. Haplogroups were assigned using Haplogrep 2.1.16 (ref. [66]) and PhyloTree mtDNA tree Build 17 (18 February 2016)[67] (Supplementary Data 4). For Y haplogroup inference we called SNPs from the International Society of Genetic Genealogy (http://isogg.org; v.11.110, 21 April 2016)) from bam files using SAMtools mpileup with option -B. We extracted sites with mapping and base quality greater than 30. Insertions and deletions, and sites showing multiple alleles, were excluded (Supplementary Data 5).

We ran kinship analysis with READ[68] within each archaeological site (minimum of three individuals; Supplementary Fig. 2). When a pair of individuals with close kinship was found, such as first-degree relationships (parent–offspring or a full sibling), we excluded the individual with fewer SNPs covered from the analyses. This resulted in the removal from the analysis of iam4 (same individual as iam5), keb8 (same individual as keb1), iam6 (first-degree relative to iam004)[6] and TAF012 (first-degree relative to TAF011)[10].

Population genomics analysis of pseudohaploid data

Data from over 300 ancient Eurasian, North African and Sub-Saharan African individuals, organized according to geography and chronology (Supplementary Data 2), were downloaded, mapped and processed though the same pipeline as used for newly generated data. The full ancient DNA dataset was merged with publicly accessible modern-day individuals sampled across the globe from the Simons Genome Diversity Project (SGDP) dataset[34] for a 2.2 million SNP panel[64]. Alleles were sampled from bam files by randomly drawing one read with MAPQ > 30 and BQ > 30 per SNP site for each ancient individual (using SAMtools v.1.5.0 mpileup with option -B), and that position was treated as (pseudo)haploid. For non-UDG-treated data (ktg001) or merged UDG and non-UDG data (iam004) we trimmed off 10 bp of sequence-ends to avoid integration of miscoding C-to-T and G-to-A substitutions. For the published partial UDG-treated data (UDG-half), 2 bp were trimmed off the sequence-ends. SNPs showing more than two alleles were excluded from the data, leaving 1,379,466 SNPs for analysis.

Principal component analysis was performed using smartpca v.10210 (ref. [69]). Principal components were calculated based on individuals from 18 Mediterranean Eurasian or North African modern-day populations from SGDP. Ancient individuals were projected onto the PCA space with options shrinkmode: YES and lsqproject: YES. An unsupervised model-based clustering algorithm, implemented in ADMIXTURE v,1.3.0 (ref. [70]), was performed for $K = 3–5$ (30 runs) on a fully pseudohaploidized, linkage disequilibrium-pruned dataset of modern-day and ancient individuals from

Mediterranean Eurasian or North African populations, leaving 812,092 SNPs for analysis. The results were parsed, aligned and plotted with pong[71].

Popstats[72] was used to calculate *f*-statistics[73], with Mbuti set as the outgroup (Supplementary Data 6 and 7). Outgroup-f_3 statistics were computed with the option –f3vanilla. Standard errors (SEs) were calculated with a weighted block jack-knife approach.

Admixture modelling was performed with qpAdm[74] using ADMIXTOOLS v.5.0, through an adapted version of qpAdm_wrapper (https://github.com/pontussk/qpAdm_wrapper) that cycles through all possible subsets of the list of source populations provided (selected based on previous results), to test one-, two-, three- and four-way admixture models. SEs were computed with 5cM block jack-knife. We used a set of 11 reference populations whose power to disentangle divergent strains of ancestry present in Europe, North Africa and the Near East has previously been described and that are differently related to the sources tested[10,31,75]. Distantly related sources were explored and, where possible, also more proximate groups (geographically, chronologically or according to standing archaeological evidence). We tried to find the most parsimonious models consistent with the data ($P > 0.05$) by checking the lowest possible number of ancestry sources necessary to explain the ancestry in each test population (Supplementary Data 8–13). The Admixture event in KTG was dated using ALDER[76] and DATES[77] (Supplementary Information 8). We calculated conditional nucleotide diversity[78] by estimation of the average number of mismatches between two individuals of the same population. SEs were estimated using a block jack-knife approach and a block size of 2,000 SNPs (Supplementary Fig. 9).

Population genomics analysis of diploid data

Diploid genotype calls for a panel of 49,791,572 SNPs were performed for northwestern African ancient individuals with at least ninefold genome coverage (oub002, ktg004 and skh001), as well as relevant, previously published ancient individuals with sequenced high-coverage genomes. Before genotype calling, base quality in read ends was reduced and indel realignment conducted with GATK 3.5.0. Diploid genotypes were called

using dbSNP v.142 as known SNPs, with GATK's UnifiedGenotyper[79]. We computed average sequencing depth (avg.DP) over all called positions for each individual and filtered for QUAL > 30 and a depth span from fivefold to 3× avg.DP per individual, using BCFtools view. This dataset was merged with data from modern-day individuals from the SGDP dataset.

Individual heterozygosity was calculated from the number of variable positions divided by that of sequenced SNPs, using the –het command in PLINK 1.9 (ref. [80]). We estimated the length and number of runs of homozygosity after filtering with the command PLINK –geno 0. MSMC[81] input files were generated from VCF files. Filters for MAPQ > 30, minimum genotype quality of 50 and sequencing depth were used. Sites not passing these filters were masked out per individual. MSMC 0.1.0 was then run for each individual.

Ethics and inclusion statement

The sampling for this study emerged from archaeology projects that involved local universities and researchers, including Y.B., whose involvement in research design included the selection of archaeological material for analyses as well as sampling supervision. The local relevance of this research is tied to the region's history, and it is locally relevant in regard to describing the human past in northwestern Africa. The study was undertaken with the highest standards of archaeogenomic research, and relevant research by local scholars was cited.

Reporting summary

Further information on research design is available in the Nature Portfolio Reporting Summary linked to this article.

Data availability

The sequence data generated for this study are available from the European Nucleotide Archive under accession no. PRJEB59008.

References

1. Linstädter, J., Medved, I., Solich, M. & Weniger, G.-C. Neolithisation process within the Alboran territory: models and possible African impact. *Quat. Int.* **274**, 219–232 (2012).

2. Martínez-Sánchez, R. M., Vera-Rodríguez, J. C., Pérez-Jordà, G., Peña-Chocarro, L. & Bokbot, Y. The beginning of the Neolithic in northwestern Morocco. *Quat. Int.* **470**, 485–496 (2018).

3. Zilhão, J. Early prehistoric navigation in the Western Mediterranean: implications for the Neolithic transition in Iberia and the Maghreb. *Isl. Archaeol. Orig. Seafaring East. Mediterr.* **11**, 185–200 (2014).

4. Mulazzani, S. et al. The emergence of the Neolithic in North Africa: a new model for the Eastern Maghreb. *Quat. Int.* **410**, 123–143 (2016).

5. Linstädter, J. The Epipalaeolithic-Neolithic-Transition in the Mediterranean region of Northwest Africa. *Quartär. International Yearbook for Ice Age and Stone Age Research* **55**, 41–62 (2008).

6. Fregel, R. et al. Ancient genomes from North Africa evidence prehistoric migrations to the Maghreb from both the Levant and Europe. *Proc. Natl Acad. Sci. USA* **115**, 6774–6779 (2018).

7. Henn, B. M. et al. Genomic ancestry of North Africans supports back-to-Africa migrations. *PLoS Genet.* **8**, e1002397 (2012).

8. Arauna, L. R. et al. Recent historical migrations have shaped the gene pool of Arabs and Berbers in North Africa. *Mol. Biol. Evol.* **34**, 318–329 (2017).

9. Hublin, J.-J. et al. New fossils from Jebel Irhoud, Morocco and the pan-African origin of *Homo sapiens*. *Nature* **546**, 289–292 (2017).

10. van de Loosdrecht, M. et al. Pleistocene North African genomes link Near Eastern and sub-Saharan African human populations. *Science* **360**, 548–552 (2018).

11. Skoglund, P. et al. Origins and genetic legacy of Neolithic farmers and hunter-gatherers in Europe. *Science* **336**, 466–469 (2012).

12. Omrak, A. et al. Genomic evidence establishes Anatolia as the source of the European Neolithic gene pool. *Curr. Biol.* **26**, 270–275 (2016).

13. Antonio, M. L. et al. Ancient Rome: a genetic crossroads of Europe and the Mediterranean. *Science* **366**, 708–714 (2019).

14. Yu, H. et al. Genomic and dietary discontinuities during the Mesolithic and Neolithic in Sicily. *iScience* **25**, 104244 (2022).

15. Günther, T. et al. Ancient genomes link early farmers from Atapuerca in Spain to modern-day Basques. *Proc. Natl Acad. Sci. USA* **112**, 11917–11922 (2015).

16. Barnett, W. K. Cardial pottery and the agricultural transition inMediterranean Europe. In *Europe's First Farmers* 93–116 (Cambridge Univ. Press, 2000).

17. Manen, C. et al. The Neolithic transition in the Western Mediterranean: a complex and non-linear diffusion process—the radiocarbon record revisited. *Radiocarbon* **61**, 531–571 (2019).

18. Natali, E. & Forgia, V. The beginning of the Neolithic in Southern Italy and Sicily. *Quat. Int.* **470**, 253–269 (2018).

19. Guilaine, J. A personal view of the neolithisation of the Western Mediterranean. *Quat. Int.* **470**, 211–225 (2018).

20. Bernabeu Auban, J. & Pardo-Gordó, S. La *impressa* en la península Ibérica: ¿Espejismo o realidad? Una reflexión a partir del binomio radiocarbono-cerámica (2020)

21. Linstädter, J., Broich, M. & Weninger, B. Defining the Early Neolithic of the Eastern Rif, Morocco–spatial distribution, chronological framework and impact of environmental changes. *Quat. Int.* **472**, 272–282 (2018).

22. Broodbank, C. & Lucarini, G. The dynamics of Mediterranean Africa, ca. 9600–1000 bc: an interpretative synthesis of knowns and unknowns. *J. Mediterr. Archaeol.* https://doi.org/10.17863/CAM.49028 (2020).

23. Martínez-Sánchez, R. M. et al. Reflections on the other side. A Southern Iberia origin for the first pottery production of Northern Morocco? *Open Archaeol.* **7**, 1054–1065 (2021).

24. Huysecom, E. et al. The emergence of pottery in Africa during the tenth millennium cal BC: new evidence from Ounjougou (Mali). *Antiquity* **83**, 905–917 (2009).

25. Garcea, E. A. A. Semi-permanent foragers in semi-arid environments of North Africa. *World Archaeol.* **38**, 197–219 (2006).

26. Dunne, J., Mercuri, A. M., Evershed, R. P., Bruni, S. & di Lernia, S. Earliest direct evidence of plant processing in prehistoric Saharan pottery. *Nature Plants* **3**, 16194 (2016).

27. Marshall, F. & Hildebrand, E. Cattle before crops: the beginnings of food production in Africa. *J. World Prehist.* **16**, 99–143 (2002).

28. Gautier, A. In *Droughts, Food and Culture: Ecological Change and Food Security in Africa's Later Prehistory* (ed. Hassan, F. A.) 195–207 (Springer US, 2002); https://doi.org/10.1007/0-306-47547-2_12.

29. David, B. & McNiven, I. J. *The Oxford Handbook of the Archaeology and Anthropology of Rock Art* (Oxford Univ. Press, 2018).

30. Smith, A. B. Origins and spread of pastoralism in Africa. *Annu. Rev. Anthropol.* **21**, 125–141 (1992).

31. Olalde, I. et al. The genomic history of the Iberian Peninsula over the past 8000 years. *Science* **363**, 1230–1234 (2019).

32. Valdiosera, C. et al. Four millennia of Iberian biomolecular prehistory illustrate the impact of prehistoric migrations at the far end of Eurasia. *Proc. Natl Acad. Sci. USA* **115**, 3428–3433 (2018).

33. Villalba-Mouco, V. et al. Survival of Late Pleistocene hunter-gatherer ancestry in the Iberian Peninsula. *Curr. Biol.* **29**, 1169–1177 (2019).

34. Mallick, S. et al. The Simons Genome Diversity Project: 300 genomes from 142 diverse populations. *Nature* **538**, 201–206 (2016).

35. Humphrey, L. T. et al. Earliest evidence for caries and exploitation of starchy plant foods in Pleistocene hunter-gatherers from Morocco. *Proc. Natl Acad. Sci. USA* **111**, 954–959 (2014).

36. Turek, J. & Vintr, J. Neolit Maghrebu ve světle nových radiokarbonových dat. *Živá Archeol.* **18**, 10–15 (2016).

37. Ringbauer, H., Novembre, J. & Steinrücken, M. Parental relatedness through time revealed by runs of homozygosity in ancient DNA. *Nat. Commun.* **12**, 5425 (2021).

38. Morales, J. et al. The introduction of South-Western Asian domesticated plants in North-Western Africa: an archaeobotanical contribution from Neolithic Morocco. *Quat. Int.* **412**, 96–109 (2016).

39. Martínez Sánchez, R. M. et al. Revisiting the Epipalaeolithic-Neolithic Transition in the extreme NW of Africa: the latest results of the chronological sequence of the Cave of Kaf Taht el-Ghar (Tétouan, Morocco). *Afr. Archaeol. Rev.* **38**, 251–274 (2021).

40. Lipson, M. et al. Parallel palaeogenomic transects reveal complex genetic history of early European farmers. *Nature* **551**, 368–372 (2017).

41. Reimer, P. J. et al. The IntCal20 Northern Hemisphere radiocarbon age calibration curve (0–55 cal kBP). *Radiocarbon* **62**, 725–757 (2020).

42. Martins, H. et al. Radiocarbon dating the beginning of the Neolithic in Iberia: new results, new problems. *J. Mediterr. Archaeol.* **28**, 105–131 (2015).

43. García Borja, P., Aura Tortosa, J. E., Bernabeu Aubán, J. & Jordá Pardo, J. F. Nuevas perspectivas sobre la neolitización en la cueva de Nerja (Málaga-España): la cerámica de la sala del vestíbulo. *Zephyrvs*

https://revistas.usal.es/uno/index.php/0514-7336/article/view/7979/8431 (2010).

44. Marcus, J. H. et al. Genetic history from the Middle Neolithic to present on the Mediterranean island of Sardinia. *Nat. Commun.* **11**, 939 (2020).

45. Mathieson, I. et al. The genomic history of southeastern Europe. *Nature* **555**, 197–203 (2018).

46. Prendergast, M. E. et al. Ancient DNA reveals a multistep spread of the first herders into sub-Saharan Africa. *Science* **365**, eaaw6275 (2019).

47. Martínez Sánchez, R. M. et al. The Middle Neolithic of Morocco's North-Western Atlantic Strip: new evidence from the El-Khil Caves (Tangier). *Afr. Archaeol. Rev.* **35**, 417–442 (2018).

48. Gilman, A. *The Later Prehistory of Tangier, Morocco* (American School of Prehistoric Research, Peabody Museum of Archaeology and Ethnology, Harvard Univ., 1975).

49. Rodríguez-Varela, R. et al. Genomic analyses of pre-European conquest human remains from the Canary Islands reveal close affinity to modern North Africans. *Curr. Biol.* **27**, 3396–3402 (2017).

50. Moots, H. M. et al. A genetic history of continuity and mobility in the Iron Age Central Mediterranean. Preprint at *bioRxiv* https://doi.org/10.1101/2022.03.13.483276 (2022).

51. Yang, D. Y., Eng, B., Waye, J. S., Dudar, J. C. & Saunders, S. R. Improved DNA extraction from ancient bones using silica-based spin columns. *Am. J. Phys. Anthropol.* **105**, 539–543 (1998).

52. Dabney, J. et al. Complete mitochondrial genome sequence of a Middle Pleistocene cave bear reconstructed from ultrashort DNA fragments. *Proc. Natl Acad. Sci. USA* **110**, 15758–15763 (2013).

53. Svensson, E. et al. Genome of Peştera Muierii skull shows high diversity and low mutational load in pre-glacial Europe. *Curr. Biol.* **31**,

2973–2983 (2021).

54. Meyer, M. & Kircher, M. Illumina sequencing library preparation for highly multiplexed target capture and sequencing. *Cold Spring Harb. Protoc.* **2010**, pdb-prot5448 (2010).

55. Günther, T. et al. Population genomics of Mesolithic Scandinavia: investigating early postglacial migration routes and high-latitude adaptation. *PLoS Biol.* **16**, e2003703 (2018).

56. Gansauge, M.-T. & Meyer, M. Single-stranded DNA library preparation for the sequencing of ancient or damaged DNA. *Nat. Protoc.* **8**, 737–748 (2013).

57. Kircher, M. Analysis of high-throughput ancient DNA sequencing data. *Methods Mol. Biol*. **840**, 197–228.

58. Schubert, M., Lindgreen, S. & Orlando, L. AdapterRemoval v2: rapid adapter trimming, identification, and read merging. *BMC Res. Notes* **9**, 88 (2016).

59. Li, H. & Durbin, R. Fast and accurate short read alignment with Burrows–Wheeler transform. *Bioinformatics* **25**, 1754–1760 (2009).

60. Li, H. et al. The sequence alignment/map format and SAMtools. *Bioinformatics* **25**, 2078–2079 (2009).

61. Fu, Q. et al. A revised timescale for human evolution based on ancient mitochondrial genomes. *Curr. Biol.* **23**, 553–559 (2013).

62. Rasmussen, M. et al. An Aboriginal Australian genome reveals separate human dispersals into Asia. *Science* **334**, 94–98 (2011).

63. Jun, G. et al. Detecting and estimating contamination of human DNA samples in sequencing and array-based genotype data. *Am. J. Hum. Genet.* **91**, 839–848 (2012).

64. Fu, Q. et al. The genetic history of ice age Europe. *Nature* **534**, 200–205 (2016).

65. Danecek, P. et al. The variant call format and VCFtools. *Bioinformatics* **27**, 2156–2158 (2011).

66. Weissensteiner, H. et al. HaploGrep 2: mitochondrial haplogroup classification in the era of high-throughput sequencing. *Nucleic Acids Res.* **44**, W58–W63 (2016).

67. Van Oven, M. PhyloTree Build 17: growing the human mitochondrial DNA tree. *Forensic Sci. Int. Genet. Suppl. Ser.* **5**, e392–e394 (2015).

68. Monroy Kuhn, J. M., Jakobsson, M. & Günther, T. Estimating genetic kin relationships in prehistoric populations. *PLoS ONE* **13**, e0195491 (2018).

69. Patterson, N., Price, A. L. & Reich, D. Population structure and eigenanalysis. *PLoS Genet.* **2**, e190 (2006).

70. Alexander, D. H., Novembre, J. & Lange, K. Fast model-based estimation of ancestry in unrelated individuals. *Genome Res.* **19**, 1655–1664 (2009).

71. Behr, A. A., Liu, K. Z., Liu-Fang, G., Nakka, P. & Ramachandran, S. Pong: fast analysis and visualization of latent clusters in population genetic data. *Bioinformatics* **32**, 2817–2823 (2016).

72. Skoglund, P. et al. Genetic evidence for two founding populations of the Americas. *Nature* **525**, 104–108 (2015).

73. Patterson, N. et al. Ancient admixture in human history. *Genetics* **192**, 1065–1093 (2012).

74. Haak, W. et al. Massive migration from the steppe was a source for Indo-European languages in Europe. *Nature* **522**, 207–211 (2015).

75. Harney, É. et al. Ancient DNA from Chalcolithic Israel reveals the role of population mixture in cultural transformation. *Nat. Commun.* **9**, 3336 (2018).

76. Loh, P.-R. et al. Inferring admixture histories of human populations using linkage disequilibrium. *Genetics* **193**, 1233–1254 (2013).

77. Narasimhan, V. M. et al. The formation of human populations in South and Central Asia. *Science* **365**, eaat7487 (2019).

78. Skoglund, P. et al. Genomic diversity and admixture differs for Stone-Age Scandinavian foragers and farmers. *Science* **344**, 747–750 (2014).

79. DePristo, M. A. et al. A framework for variation discovery and genotyping using next-generation DNA sequencing data. *Nat. Genet.* **43**, 491–498 (2011).

80. Purcell, S. et al. PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am. J. Hum. Genet.* **81**, 559–575 (2007).

81. Schiffels, S. & Durbin, R. Inferring human population size and separation history from multiple genome sequences. *Nat. Genet.* **46**, 919–925 (2014).

Acknowledgements

We thank A. R. Munters for bioinformatics support and R. P. Stjerna and F. Sánchez-Quinto for helpful discussions on data analysis. Sequencing was performed at The National Genomics Infrastructure, Uppsala. Computations and data handling were enabled by resources provided by the Swedish National Infrastructure for Computing at the Uppsala Multidisciplinary Center for Advanced Computational Science, partially funded by the Swedish Research Council through grant agreement no. 2018-05973. This project was supported by the Knut and Alice Wallenberg Foundation (to M.J.), Vetenskapsrådet (grant nos. 2018-05537 and 2022-04642 to M.J. and 2017-05267 to T.G.) and Ramón y Cajal (grant no. RYC2018-025223-I to C.V.). The Spanish–Moroccan archaeological team was supported by the European Research Council (no. ERC AdG 230561).

Funding

Open access funding provided by Uppsala University.

Author information

Authors and Affiliations

1. Human Evolution, Department of Organismal Biology and SciLifeLab, Uppsala University, Uppsala, Sweden

   Luciana G. Simões, Torsten Günther & Mattias Jakobsson

2. Departamento de Historia, Universidad de Córdoba, Cordoba, Spain

   Rafael M. Martínez-Sánchez

3. Área de Prehistoria, Departamento de Historia, Geografía y Antropología, Centro de Investigación en Patrimonio Histórico, Cultural y Natural, Facultad de Humanidades, Universidad de Huelva, Huelva, Spain

   Juan Carlos Vera-Rodríguez

4. Universidad de Burgos, Departamento de Historia, Geografía y Comunicaciones, Burgos, Spain

   Eneko Iriarte & Cristina Valdiosera

5. Centre for Palaeogenetics, Stockholm, Sweden

   Ricardo Rodríguez-Varela

6. Department of Archaeology and Classical Studies, Stockholm University, Stockholm, Sweden

   Ricardo Rodríguez-Varela

7. Institut National des Sciences de l'Archéologie et du Patrimoine, Rabat, Morocco

Youssef Bokbot

8. Department of History and Archaeology, La Trobe University, Melbourne, Victoria, Australia

   Cristina Valdiosera

9. Palaeo-Research Institute, University of Johannesburg, Auckland Park, South Africa

   Mattias Jakobsson

Contributions

C.V., J.C.V.-R. and M.J. conceived the study. R.M.-S., J.C.V.-R., E.I., R.R.-V., Y.B. and C.V. selected and sampled archaeological material. L.G.S. performed DNA laboratory work. R.M.-S., J.C.V.-R. and C.V. provided archaeological interpretations. L.G.S., T.G. and M.J. analysed genetic data. L.G.S., T.G., R.M.-S., J.C.V.-R., C.V. and M.J. wrote the paper, with input from all authors.

Corresponding authors

Correspondence to Cristina Valdiosera or Mattias Jakobsson.

Ethics declarations

Competing interests

The authors declare no competing interests.

Peer review

Peer review information

Nature thanks the anonymous reviewers for their contribution to the peer review of this work.

Additional information

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Supplementary information

Supplementary Information

A description of the northwestern African archaeological sites studied and a detailed description of the methods used for analyses, supplementary results and supplementary discussions for these results.

Reporting Summary

Supplementary Data

Supplementary Data 1–15.

Rights and permissions

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by/4.0/.

[Reprints and Permissions]

About this article

Cite this article

Simões, L.G., Günther, T., Martínez-Sánchez, R.M. *et al.* Northwest African Neolithic initiated by migrants from Iberia and Levant. *Nature* **618**, 550–556 (2023). https://doi.org/10.1038/s41586-023-06166-6

- Received: 28 October 2022
- Accepted: 03 May 2023
- Published: 07 June 2023
- Issue Date: 15 June 2023
- DOI: https://doi.org/10.1038/s41586-023-06166-6

Share this article

Anyone you share the following link with will be able to read this content:

Get shareable link

Sorry, a shareable link is not currently available for this article.

Copy to clipboard

Provided by the Springer Nature SharedIt content-sharing initiative

This article is cited by

- **[Ancient DNA reveals how farming spread into northwest Africa]**

  - Louise Humphrey
  - Abdeljalil Bouzouggar

Nature (2023)

Ancient DNA reveals how farming spread into northwest Africa

- Louise Humphrey
- Abdeljalil Bouzouggar

Nature News & Views 07 Jun 2023

This article was downloaded by **calibre** from https://www.nature.com/articles/s41586-023-06166-6

- Article
- Open Access
- Published: 31 May 2023

Geometric constraints on human brain function

- James C. Pang ,
- Kevin M. Aquino,
- Marianne Oldehinkel,
- Peter A. Robinson,
- Ben D. Fulcher,
- Michael Breakspear &
- …
- Alex Fornito

Show authors

Nature **volume 618**, pages 566–574 (2023)

- 53k Accesses
- 1 Citations
- 902 Altmetric
- Metrics details

Abstract

The anatomy of the brain necessarily constrains its function, but precisely how remains unclear. The classical and dominant paradigm in neuroscience is that neuronal dynamics are driven by interactions between discrete,

functionally specialized cell populations connected by a complex array of axonal fibres[1,2,3]. However, predictions from neural field theory, an established mathematical framework for modelling large-scale brain activity[4,5,6], suggest that the geometry of the brain may represent a more fundamental constraint on dynamics than complex interregional connectivity[7,8]. Here, we confirm these theoretical predictions by analysing human magnetic resonance imaging data acquired under spontaneous and diverse task-evoked conditions. Specifically, we show that cortical and subcortical activity can be parsimoniously understood as resulting from excitations of fundamental, resonant modes of the brain's geometry (that is, its shape) rather than from modes of complex interregional connectivity, as classically assumed. We then use these geometric modes to show that task-evoked activations across over 10,000 brain maps are not confined to focal areas, as widely believed, but instead excite brain-wide modes with wavelengths spanning over 60 mm. Finally, we confirm predictions that the close link between geometry and function is explained by a dominant role for wave-like activity, showing that wave dynamics can reproduce numerous canonical spatiotemporal properties of spontaneous and evoked recordings. Our findings challenge prevailing views and identify a previously underappreciated role of geometry in shaping function, as predicted by a unifying and physically principled model of brain-wide dynamics.

Main

The dynamics of many natural systems are fundamentally constrained by their underlying structure. For instance, the shape of a drum influences its acoustic properties, the morphology of a river bed shapes underwater currents and the geometry of a protein determines the molecules with which it can interact[9]. The nervous system is no exception, with the rich and complex spatiotemporal dynamics of anatomically distributed neuronal populations being supported by their intricate web of axonal interconnectivity[1,10]. Several studies have shown correlations between various properties of brain connectivity and activity[11], but precisely how spatiotemporal patterns of neural dynamics are constrained by a relatively stable neuroanatomical scaffold remains unclear.

In diverse areas of physics and engineering, structural constraints on system dynamics can be understood via the system's eigenmodes, which are fundamental spatial patterns corresponding to the natural, resonant modes of the system[12]. In the linear regime, such as brain activity under normal (that is, non-seizure-like) conditions[13], eigenmodes (hereafter also referred to as modes) offer a particularly powerful and rigorous formalism for linking brain anatomy with the physical processes that shape activity. Through this lens, spatiotemporal patterns of neuronal dynamics emerge from excitations of the brain's structural eigenmodes, much like the harmonics of a plucked violin string arise from vibrations of its own resonant modes.

Critically, just as the resonant frequencies of a violin string are determined by its length, density and tension, the eigenmodes of the brain are determined by its structural—physical, geometric and anatomical—properties. Do any of these specific structural properties make a dominant contribution to dynamics? Here we test two influential and competing theories that make different predictions about which key elements of brain structure shape function.

One classical perspective, which represents the dominant paradigm in neuroscience, has its roots in Ramon y Cajal's neuron doctrine[14], Brodmann's cytoarchitectonics[15] and over a century of work localizing functions to specific brain regions[16]. According to this view, spatiotemporal patterns of neural dynamics arise from interactions between discrete, functionally specialized cell populations connected by a topologically complex array of short- and long-range axonal connections[2,3]. In humans, these connections can be estimated at macroscopic scales by diffusion magnetic resonance imaging (dMRI) to yield a structural connectivity matrix, called a connectome[17]. This approach has been used extensively to understand brain organization and dynamics[2,17,18], and recent work has proposed that eigenmodes derived from such a discrete connectome—referred to here as connectome eigenmodes—can be used to reconstruct the spatial patterns of canonical functional networks of the human cortex mapped with functional MRI (fMRI)[19,20].

One limitation of this discrete connectome-based view is that it relies on an abstract representation of brain anatomy that does not directly account for its

physical properties and spatial embedding (that is, geometry and topology). These characteristics are explicitly incorporated into a broad class of neural field theories (NFTs)[4,5,6] that describe mean-field neural dynamics on spatial scales above 0.5 mm (Supplementary Information 1). One physiologically constrained form of NFT has unified a diverse range of empirical phenomena[6,21] by treating cortical activity as a superposition of travelling waves propagating through a physically continuous sheet of neural tissue. In this theory, neural interactions between different cortical locations are approximated by a homogeneous spatial kernel that declines roughly exponentially with distance[22], in line with experimental evidence that the organization of the nervous systems of numerous species is universally governed by an exponential distance rule (EDR) for connectivity[10,23,24].

Given wave-like dynamics and EDR-like connectivity, a key prediction of NFT is that the intrinsic geometry of the brain physically shapes and imposes boundary conditions on emergent dynamics[7,8,25]. A remarkable corollary of this view is that, if we prioritize spatial and physical constraints on brain anatomy, we only need to consider the shape of the brain, and not its full array of topologically complex axonal interconnectivity, to understand spatially patterned activity. More formally, the theory predicts that eigenmodes derived from brain geometry—hereafter referred to as geometric eigenmodes—represent a more fundamental anatomical constraint on dynamics than the connectome[7,8,25]. This view stands in stark contrast to the classical view that complex patterns of interregional anatomical connectivity shape brain activity[26].

Here we test these competing views of the brain with the aim of identifying the principal structural constraints on human brain dynamics. In line with theoretical predictions from NFT, we show that diverse experimental fMRI data from spontaneous and task-evoked recordings in the human neocortex can be explained more parsimoniously by eigenmodes derived from cortical geometry (geometric eigenmodes) than by those obtained from measures of brain connectivity (connectome eigenmodes). We further confirm that stimulus-evoked activity is dominated by excitations of geometric eigenmodes with long spatial wavelengths, challenging classical views that such activity is localized to focal, spatially isolated clusters. To directly link these structural constraints to the physical processes driving brain dynamics,

we use a generative model to show how wave dynamics unfolding on the geometry of the cortex can explain diverse features of functional brain organization. Finally, we show that the close relationship between geometry and function captured by eigenmodes extends to non-neocortical structures, indicating that this link is a universal property of brain organization.

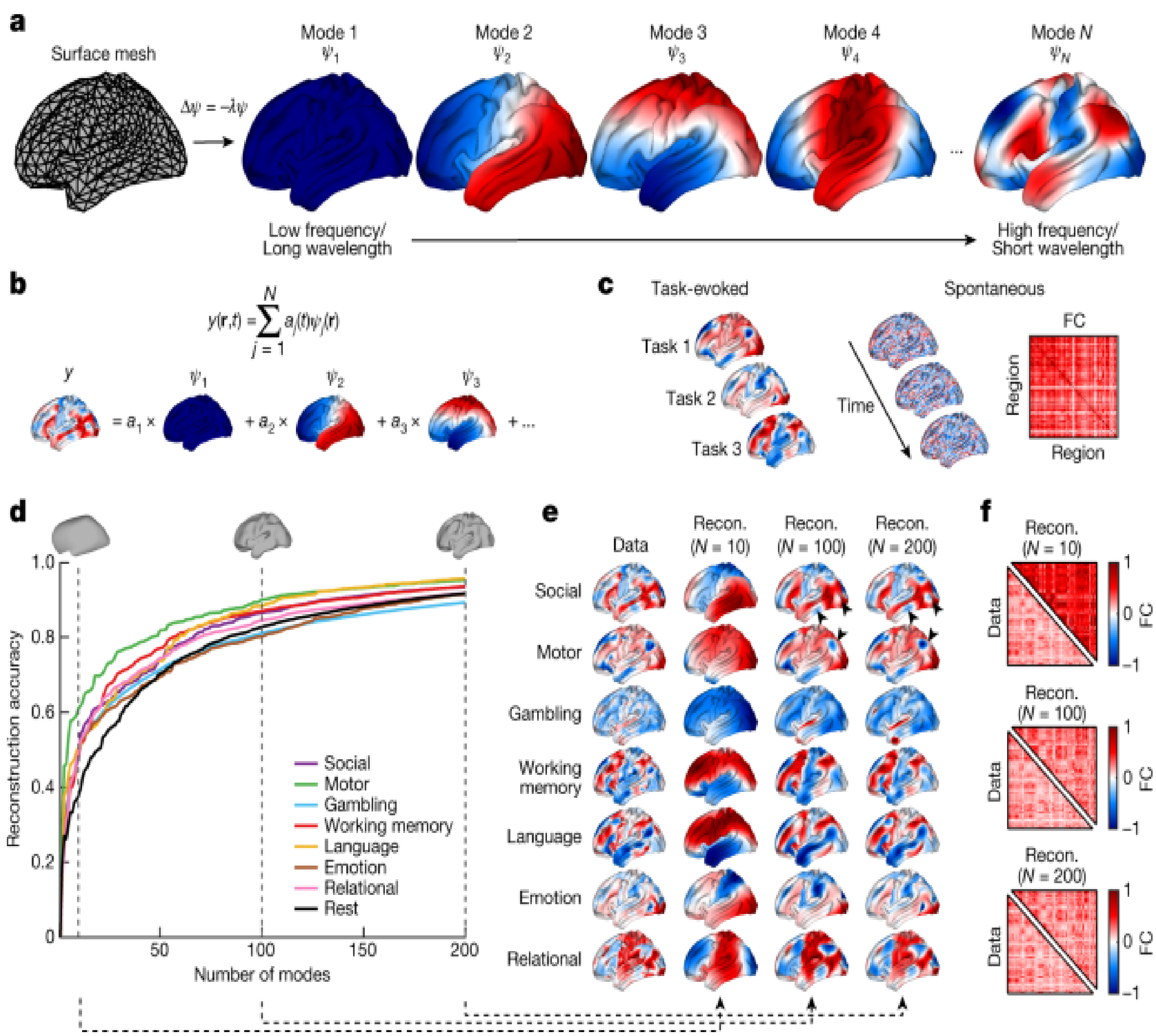
Geometric modes constrain cortical activity

We first examine the degree to which geometric eigenmodes can explain diverse aspects of human neocortical activity. To derive the eigenmodes, we use a mesh representation of a population-averaged template of the neocortical surface (Fig. 1a and Derivation of cortical geometric eigenmodes in Methods). We then construct the Laplace–Beltrami operator (LBO) from this surface mesh, which captures local vertex-to-vertex spatial relations and curvature, and solve the eigenvalue problem,

$${\nabla }^{2}\psi =\Delta \psi =-\,\lambda \psi ,$$
(1)

where ∇ is the gradient operator, Δ is the LBO and $\psi =\{{\psi }_{1}({\bf{r}}),{\psi }_{2}({\bf{r}}),\ldots \}$ is the family of geometric eigenmodes with the corresponding family of eigenvalues $\lambda =\{{\lambda }_{1},{\lambda }_{2},\ldots \}$. The eigenvalues are ordered sequentially according to the spatial frequency or wavelength of the spatial patterns of each mode (Fig. 1a and Extended Data Fig. 1), such that ψ_1 is the mode with the longest wavelength. The resulting eigenmodes are orthogonal, forming a complete basis set to decompose spatiotemporal dynamics unfolding on the cortex as a weighted sum of modes with varying wavelengths (Fig. 1b and Modal decomposition of brain activity in Methods). Unless otherwise specified, we use $N = 200$ modes throughout this study.

Fig. 1: Reconstruction of neocortical activity with geometric eigenmodes.



a, Geometric eigenmodes are derived from the cortical surface mesh by solving the eigenvalue problem, $\Delta \psi =-\,\lambda \psi$ (equation (1)). The modes ${\psi }_{1},\,{\psi }_{2},{\psi }_{3},\ldots ,\,{\psi }_{N}$ are ordered from low to high spatial frequency (long to short spatial wavelengths). Negative, zero and positive values are coloured blue, white and red, respectively. **b**, Modal decomposition of brain activity data. The example shows how a spatial map, $y({\bf{r}}\,,t)$, at a given time, *t*, can be decomposed as a sum of modes, ψ_j, weighted by a_j. **c**, Left, we reconstruct task-evoked data using spatial maps of activation for a diverse range of stimulus contrasts. Right, we reconstruct spontaneous activity by decomposing the spatial map at each time frame and generating a region-to-region FC matrix. **d**, Reconstruction accuracy of seven key HCP task-contrast maps (Supplementary Information 2.1) and resting-state FC as a function of the number of modes. Insets show cortical surface

reconstructions, demonstrating the spatial scales relevant to the first 10, 100 and 200 modes corresponding to spatial wavelengths of approximately 120, 40 and 30 mm, respectively. **e**, Group-averaged empirical task-activation maps and reconstructions (recon.) obtained using 10, 100 and 200 modes of the seven key HCP task contrasts. Black arrowheads indicate localized activation patterns that are more accurately reconstructed when using short-wavelength modes. **f**, Group-averaged empirical resting-state FC matrices and reconstructions using 10, 100 and 200 modes.

Using this decomposition we evaluate the accuracy of geometric eigenmodes in capturing both task-evoked and spontaneous brain activity (Fig. 1c) measured in 255 healthy individuals from the Human Connectome Project[27] (HCP; HCP data in Methods and Supplementary Information 2). For task-evoked activity, we map 47 task-based contrasts drawn from seven different tasks representing distinct evoked activation patterns. We then reconstruct each individual's activation map using an increasing number of modes up to a maximum of 200 (Fig. 1d). For spontaneous, task-free (so-called resting-state) activity, we reconstruct the spatial map of activity at each time frame and then generate a region-to-region functional coupling (FC) matrix, describing correlations of activity among 180 discrete brain regions per hemisphere[28]. To allow direct comparison between task-evoked and spontaneous recordings, we apply the same regional parcellation to the task-evoked data (Cortical parcellations in Methods). Finally, we quantify reconstruction accuracy by calculating the correlation between empirical and reconstructed task-evoked activation maps and spontaneous FC matrices (Fig. 1d–f).

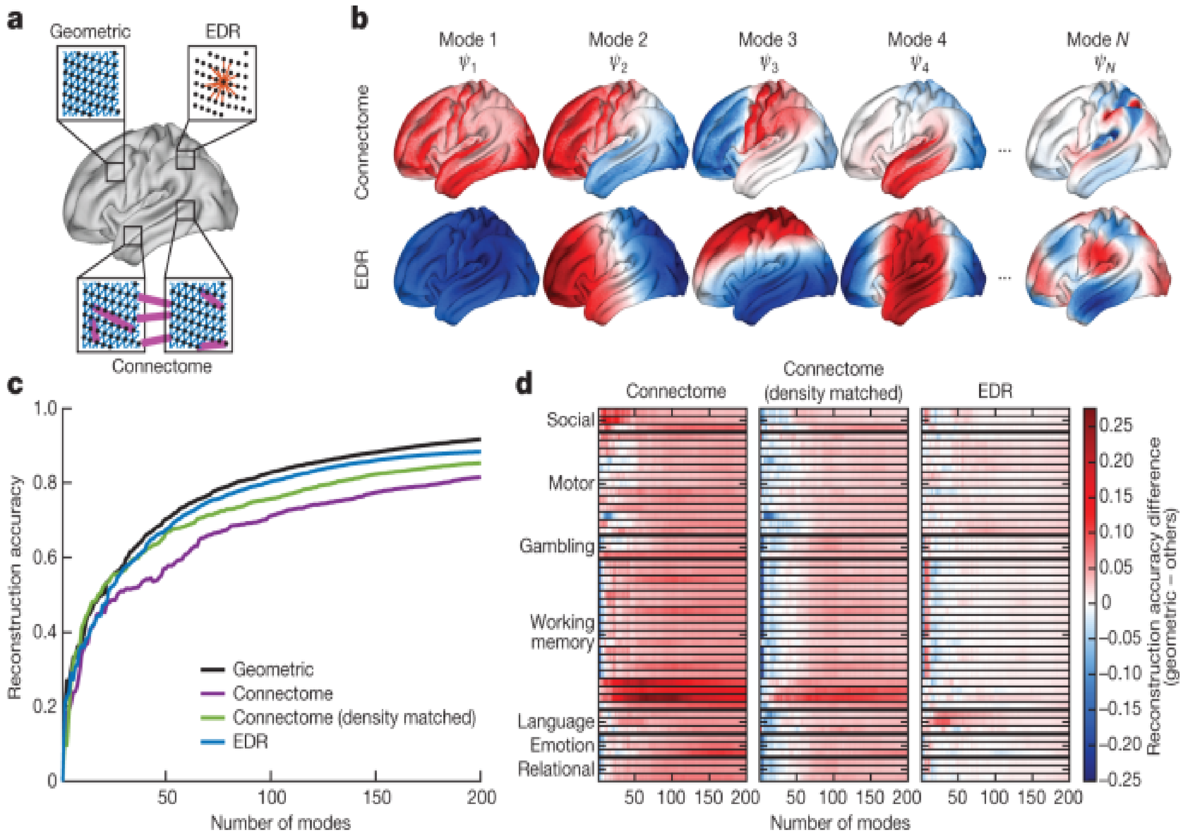
We observe that reconstruction accuracy increases with an increasing number of modes across all task contrasts and in the resting state, with $r \geq 0.38$ already achieved using just $N = 10$ modes (Fig. 1d). Large-scale modes are also differentially recruited across different tasks, suggesting that particular stimuli excite specific modes (Fig. 1e). Improvements in reconstruction accuracy become slow after ten modes, reaching $r \geq 0.80$ at approximately $N = 100$ modes, with only incremental increases in reconstruction accuracy beyond this point. Beause the first 100 modes have wavelengths above around 40 mm (Supplementary Table 1), and the inclusion of shorter-wavelength modes only refines reconstruction of localized patterns (arrowheads in Fig. 1e), our findings suggest that the data

are predominantly comprised of spatial patterns with long spatial wavelengths (see next section for a more detailed analysis).

These results are consistent across all 47 HCP task contrasts (Supplementary Fig. 1) and parcellations of varying resolutions (Supplementary Fig. 2), but data parcellated at higher resolution require more modes to achieve high reconstruction accuracy due to the low-pass spatial filtering effect of coarser parcellations. Our results are also not affected by the use of a population-averaged cortical surface template (rather than individual-specific surfaces) to derive the geometric eigenmodes (Supplementary Figs. 3–5 and Supplementary Information 3). Together these findings indicate that cortical geometric eigenmodes form a compact representation that captures diverse aspects of task-evoked and spontaneous cortical activity. Moreover, they show that such activity is dominated by long-wavelength, large-scale eigenmodes.

We next test the hypothesis that geometric eigenmodes provide a more parsimonious and fundamental description of dynamics than eigenmodes derived from a graph-based connectome approximation. To this end we compare the reconstruction accuracy of geometric eigenmodes against three alternative connectome-derived eigenmode basis sets (see Fig. 2a for a schematic). The first basis set is derived empirically from a connectome mapped with dMRI tractography at vertex resolution and thresholded to obtain a connection density of 0.10%, as done previously[29] (Derivation of connectome eigenmodes in Methods). The second basis set is derived from a connectome constructed synthetically according to a homogeneous stochastic wiring process governed by an exponential distance-dependent connection probability to mimic simple, EDR-like connectivity (Derivation of EDR eigenmodes in Methods). Because the connection densities of empirical and EDR connectomes differed, we evaluated a third basis set derived from the empirical connectome thresholded at 1.55% to match the density of the EDR connectome. The connectome, EDR and density-matched connectome eigenmodes described above are derived from the graph Laplacian (a discrete counterpart of the LBO) of their respective connectivity matrices (Fig. 2b and Extended Data Fig. 1).

Fig. 2: Geometric eigenmodes benchmarked against connectome-based eigenmodes.



a, Schematic of the anatomical properties used to derive eigenmodes for cortical geometry, the connectome and the EDR connectome. Geometric eigenmodes rely on local surface mesh information such as links (blue) between neighbouring surface mesh vertices (dots) and curvature. Connectome eigenmodes rely on local links between mesh vertices (blue) and short- and long-range connections (magenta) reconstructed empirically from dMRI. EDR eigenmodes rely on connections (red) generated from a stochastic wiring process in which the probability of connection between vertices exponentially decays as a function of their distance. **b**, Example connectome and EDR eigenmodes. Negative, zero and positive values are coloured blue, white and red, respectively. Despite some similarities, the spatial patterns of the modes are distinct from those derived using cortical geometry (compare with Fig. 1a). **c**, Reconstruction accuracy of resting-state FC matrices achieved by geometric, EDR and two variants of connectome eigenmodes: one using a connectome as defined using previous methods[29] and the other with the same connection density as the EDR connectome to allow fair comparison (for other densities see Supplementary Figs. 6 and 7). **d**, Difference in reconstruction accuracy of all 47 HCP task-contrast maps

achieved by geometric eigenmodes and the other basis sets, as indicated by the text above each panel. Each row represents a different task contrast, grouped here by broad types (Supplementary Information 2.1); red indicates superior performance for geometric eigenmodes. Note that while there seems to be a performance advantage for connectome eigenmodes for reconstructions incorporating fewer than ten modes relative to geometric eigenmodes, reconstruction accuracy is generally low (average $r = 0.42$ across the different tasks) compared with that for 100 modes (average $r = 0.71$).

To summarize, geometric eigenmodes account for the intrinsic curvature of the cortical surface and local vertex-to-vertex relations in the surface mesh; connectome eigenmodes do not consider curvature but capture local spatial relations between mesh vertices, along with short- and long-range connections measured with dMRI; and EDR eigenmodes account for the effect of a homogeneous, stochastic, distance-dependent connection rule without fully capturing the cortical geometry (Fig. 2a). Contrasting these different basis sets thus allows us to disentangle the contributions to brain dynamics of cortical geometry from structural connectivity.

Direct comparison of the reconstruction accuracy of these different basis sets shows that geometric eigenmodes consistently show the highest reconstruction accuracy across both spontaneous (Fig. 2c) and task-evoked (Fig. 2d) data. EDR eigenmodes perform nearly as well as geometric eigenmodes whereas connectome eigenmodes are the least accurate. This finding holds true regardless of the parcellation used (Extended Data Figs. 2 and 3), the specific connection density used to generate the connectome eigenmodes (Supplementary Figs. 6 and 7 and Supplementary Information 4) and whether we generate the connectome using a discrete regional parcellation rather than at vertex resolution (Supplementary Fig. 8 and Supplementary Information 4). We additionally find that geometric eigenmodes show stronger out-of-sample generalization than principal components of the functional data themselves (calculated via principal component analysis (PCA); Supplementary Fig. 9, Extended Data Fig. 4 and Supplementary Information 5) and better performance than Fourier spatial basis sets (Extended Data Fig. 5, Supplementary Information 6 and Comparisons with statistical basis sets in Methods).

Taken together, these results demonstrate the parsimony, robustness and generality of geometric eigenmodes as a basis set for brain function. They also support the prediction of NFT that brain activity is best represented in terms of eigenmodes derived directly from the shape of the cortex, thus emphasizing a fundamental role of geometry in constraining dynamics.

Long wavelengths dominate cortical activity

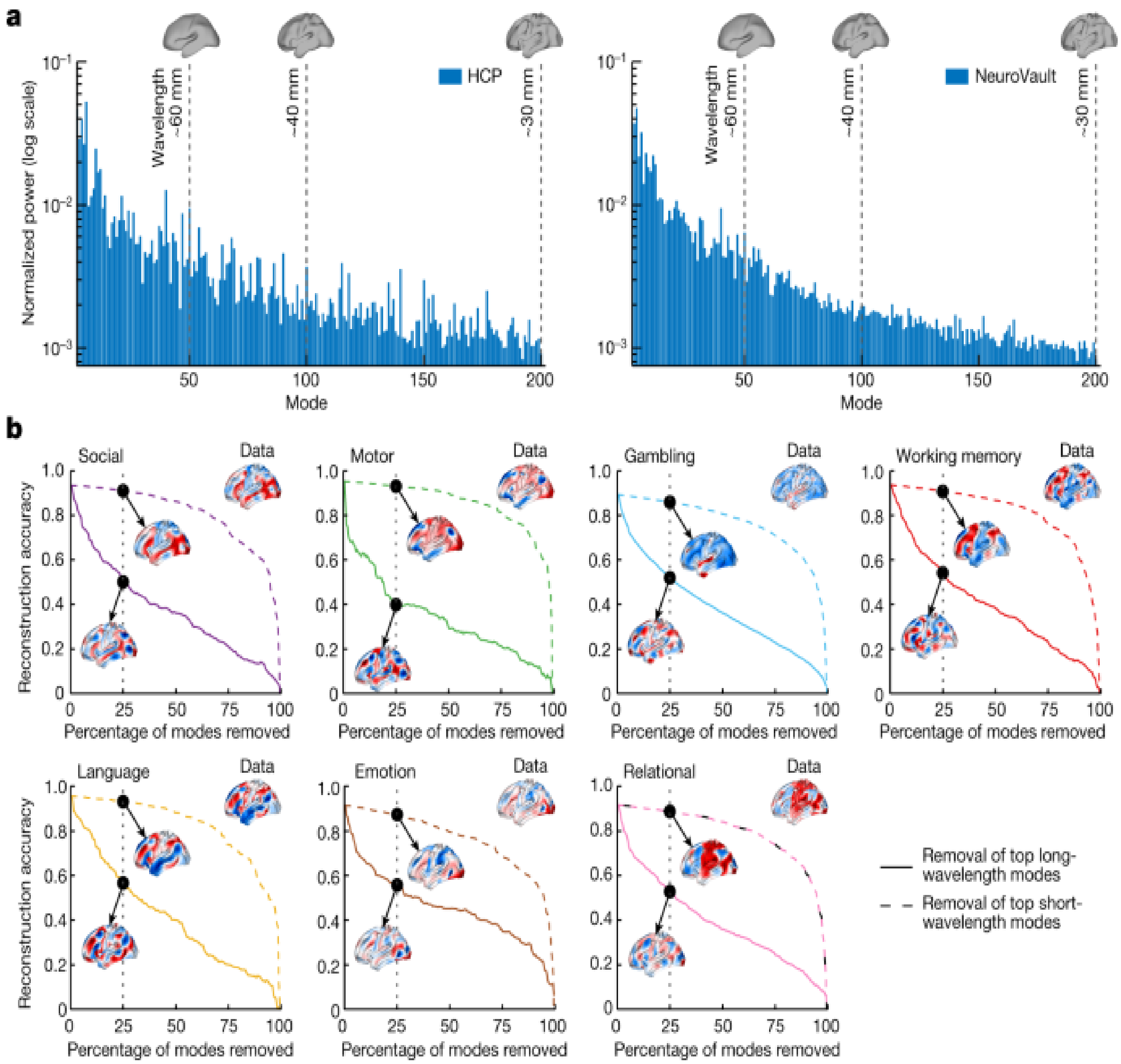
Reconstructions of both spontaneous and task-evoked data with geometric eigenmodes show that the spatial organization of brain activity is dominated by patterns with spatial wavelengths of about 40 mm or longer (Fig. 1d–f). This result counters classical approaches to analysis of neuroimaging data, in which stimulus-evoked activations are mapped by thresholding statistical maps to identify focal, isolated areas of heightened activity. This classical approach rests on the assumption that focal loci represent discrete brain regions putatively engaged by the stimulus and that subthreshold activity in other regions is of negligible interest. The surprisingly long-wavelength content of task-activation data (Fig. 1d–e) suggests that classical procedures focus only on the tips of the iceberg and obscure the underlying spatially extended and structured patterns of activity evoked by the task (see Extended Data Fig. 6 for an explanation of the reasons involved). These observations accord with the theoretical predictions of NFT and previous analyses of task-evoked electroencephalography (EEG) signals[30,31].

Here we leverage the modal decomposition described in Fig. 1b to characterize the complete spatial pattern—the entire iceberg—of task-evoked activation. To this end we analyse the spatial power spectrum obtained using a geometric mode decomposition of group-averaged unthresholded activation maps from the 47 task contrasts in HCP[27,32] (Modal power spectra of task-evoked activation maps in Methods). As an independent replication, we also analyse 10,000 unthresholded activation maps from 1,178 independent experiments available in the NeuroVault repository[33], thus providing a comprehensive picture of the diversity of stimulus-evoked activation patterns mapped in the human brain.

Despite the wide range of stimuli, paradigms and data-processing approaches used to acquire these activation maps, we observe that a large

fraction of power in the maps is concentrated in the first 50 modes, corresponding to spatial wavelengths greater than around 60 mm (Fig. 3a; similar results are found separately for each of the key HCP task-contrast maps; Extended Data Fig. 7). Using surrogate data, we confirm that these findings cannot be explained by the spatial smoothing induced by typical fMRI processing pipelines, which can filter out short-wavelength spatial patterns of activity (Extended Data Fig. 8 and Supplementary Information 7). We further observe that incremental, sequential removal of long-wavelength modes has a much greater impact on reconstruction accuracy than removal of short-wavelength modes (Fig. 3b and Contributions of long- and short-wavelength modes in Methods). For instance, across the seven key HCP task contrasts, removal of the top 25% long-wavelength modes (modes 1–50) yields a drop in reconstruction accuracy of around 40–60% whereas removal of the top 25% short-wavelength modes (modes 151–200) yields a drop of only around 2–4% (Fig. 3b, insets). These results indicate that, on temporal and spatial scales accessible with fMRI, evoked cortical activity comprises large-scale, nearly brain-wide spatial patterns, challenging classical views that such activity should be described in terms of discrete, isolated and anatomically localized activation clusters.

Fig. 3: Task-evoked activity excites long-wavelength modes.



a, Normalized mean power spectra of the 47 HCP task-contrast maps (left) and 10,000 contrast maps from the NeuroVault database (right). Insets show cortical surface reconstructions demonstrating spatial scales relevant to the first 50, 100 and 200 modes corresponding to spatial wavelengths of approximately 60, 40 and 30 mm, respectively. Contrast-specific spectra for the seven key HCP task contrasts are presented in Extended Data Fig. 7. **b**, Reconstruction accuracy of the seven key HCP task-contrast maps as a function of the percentage of modes (among 200) removed in the reconstruction process. Solid and dashed lines correspond to removal of the top long- and short-wavelength modes, respectively. Insets show group-averaged empirical activation maps (data) and their reconstructions after

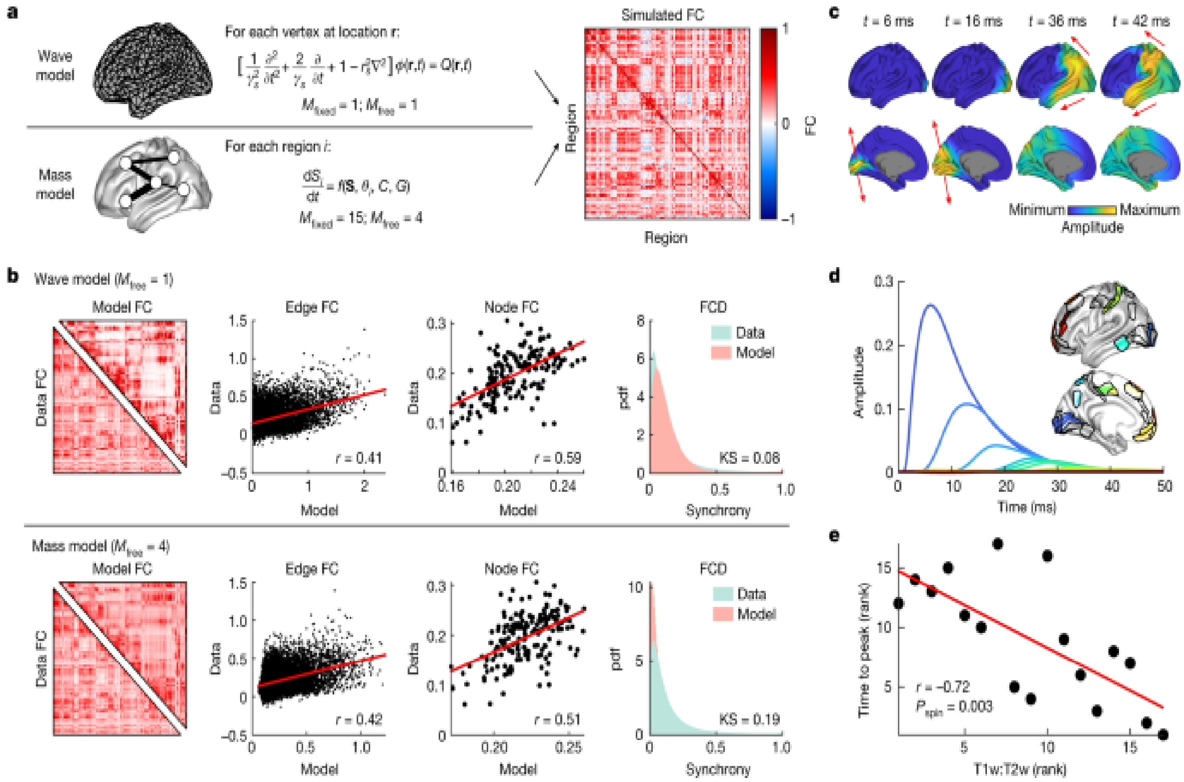
removal of 25% of modes. Negative, zero and positive values are coloured blue, white and red, respectively.

Wave dynamics bridge geometry and function

Geometric eigenmodes of the cortex are obtained by solving the eigenvalue problem of the LBO, which is also known as the Helmholtz equation (equation (1)). In physically continuous systems, the solutions of the Helmholtz equation correspond to the spatial projections of the solutions of a more general wave equation, such that the resulting eigenmodes inherently represent the vibrational patterns, or standing waves, of the system's dynamics[34]. This equivalence implies that the superior efficacy of geometric eigenmodes in the reconstruction of diverse patterns of brain activity results from a fundamental role of wave dynamics in shaping these patterns, as predicted by NFT. This prediction has been confirmed through models of EEG recordings[21,35], but waves across the whole brain have only recently been observed in fMRI signals[36,37] and thus far lack a theoretical explanation. Here we use NFT and geometric eigenmodes to show that wave dynamics can provide a unifying account of diverse empirical and physiological phenomena observed at scales accessible with fMRI.

We model neural activity using an isotropic damped NFT wave equation without regeneration[6] (Fig. 4a and NFT wave model in Methods). Under this model, activity propagates between points on the neocortex through their white-matter connectivity with a strength that decays approximately exponentially with distance (Supplementary Fig. 10 and Supplementary Information 1 and 8). To simulate resting-state neural activity we use a white noise input to mimic unstructured stochastic fluctuations[21] (Modelling resting-state dynamics in Methods). We compare the performance of this simple wave model with a biophysically based neural mass model (balanced excitation–inhibition (BEI) model) that has been used extensively to understand resting-state fMRI signals[38] (Fig. 4a and Neural mass model in Methods). The neural mass model is closely aligned with the classical, connectome-centric view of brain function, representing dynamics as the result of interactions between neuronal populations in discrete anatomical regions, coupled according to an empirically measured connectome.

Fig. 4: Wave dynamics shape patterns of spontaneous and stimulus-evoked activity.



a, For the wave model, the activity $\varphi ({\bf{r}},t)$ at location **r** and time *t* is governed by a wave equation with damping rate γ_s, spatial length scale r_s and input Q. For the neural mass model the activity ${S}_{i}\left(t\right)$ of region *i* is described by the function *f*, which depends on the activity of other regions *S*, local population parameters θ_i and connectome *C* scaled by the global coupling parameter *G*. The model dynamics are used to calculate a simulated FC matrix (Methods). M_{fixed} and M_{free} correspond to the number of fixed and free parameters of each model, respectively. **b**, Comparison of data and model simulations based on various metrics which, from left to right, are: FC matrix (for visual purposes), edge FC, node FC and FCD. For edge FC and node FC, red lines represent linear fits with Pearson correlation coefficient *r*; for FCD, the probability density function (pdf) of the similarity of interregional synchrony in data and model dynamics is compared using the Kolmogorov–Smirnov (KS) statistic. **c**, Wave propagation of activity after 1 ms stimulation of V1 for t = 1–2 ms. Arrows indicate the direction of propagation (Supplementary Video 1). **d**, Activity profile of 17 regions in the visual cortical hierarchy. Insets show

spatial locations of regions on the cortical surface coloured according to their activation profiles. **e**, Relationship of ranked activity profile time to peak and ranked T1w:T2w value of the regions in **d**. Red line represents a linear fit of the ranked variables, with Spearman correlation coefficient *r* and one-sided P_{spin}, estimated from 10,000 permutations.

We first compare the efficacy of the two models in capturing distinct and commonly studied properties of spontaneous, task-free FC: namely, static pairwise FC (edge FC), static node-level average FC (node FC) and time-resolved dynamic properties of FC (FCD) (Modelling resting-state dynamics in Methods). Across all FC-based benchmark measures, the wave model shows comparable or superior performance in reconstruction of empirical data relative to the neural mass model (Fig. 4b). The wave model also captures time-lagged properties[36,37,39] of empirical resting-state activity more accurately than the mass model (Extended Data Fig. 9 and Measuring time-lagged properties of resting-state dynamics in Methods). This strong performance of the wave model is remarkable given its relative simplicity: the wave model only requires the geometry of the cortex (that is, the surface mesh) as input and includes one fixed parameter and one free parameter (r_s) for fitting to data (Extended Data Fig. 10) whereas the neural mass model requires a dMRI-derived interregional anatomical connectivity matrix and comprises 15 fixed parameters and four free parameters (Supplementary Information 9). These considerations indicate that wave dynamics offer a more accurate and parsimonious mechanistic account of macroscale, spontaneous cortical dynamics captured by fMRI.

We next consider stimulus-evoked cortical activity in the wave model. We analyse cortical responses to sensory stimulation of primary visual cortex (V1), because it elicits a well-defined hierarchy of regional cortical responses[40,41] (Modelling stimulus-evoked dynamics in Methods). A 1 ms pulse input to V1 yields a propagating wave of activity that rapidly splits along the dorsal and ventral visual processing streams (Fig. 4c (arrows) and Supplementary Video 1), consistent with the mainstream understanding of hierarchical visual processing[42]. Remarkably, this result indicates that geometric constraints on travelling waves of evoked activity are sufficient for the segregation of the dorsal and ventral processing streams, which have traditionally been thought to be driven primarily by complex patterns of

layer-specific connectivity[40,42,43]. Furthermore, the temporal profile of evoked responses across the visual system follows a well-defined timescale hierarchy, with higher-order association areas showing peak responses that are delayed and prolonged compared with lower-order visual areas (Fig. 4d). These findings thus indicate that this hierarchical ordering, previously identified in experimental and modelling studies[41,44,45], emerges naturally from waves of excitation propagating through the cortical medium. Critically, this hierarchical temporal ordering of areal responses strongly correlates with an independent anatomical measure of the cortical processing hierarchy based on non-invasive estimates of myeloarchitecture (T1-weighted (T1w) and T2-weighted (T2w) ratio)[46,47]. This correlation is particularly strong within the visual processing hierarchy ($r = -0.72$, one-sided spin-test P value (P_{spin}) = 0.003; Fig. 4e) but is also present when considering all cortical areas ($r = -0.44$, $P_{spin} = 0.037$; Supplementary Fig. 11). Together, our modelling results show how simple wave dynamics unfolding on the geometry of the cortex provide a unifying generative mechanism for capturing complex properties of spatiotemporal brain activity.

Geometry constrains subcortical activity

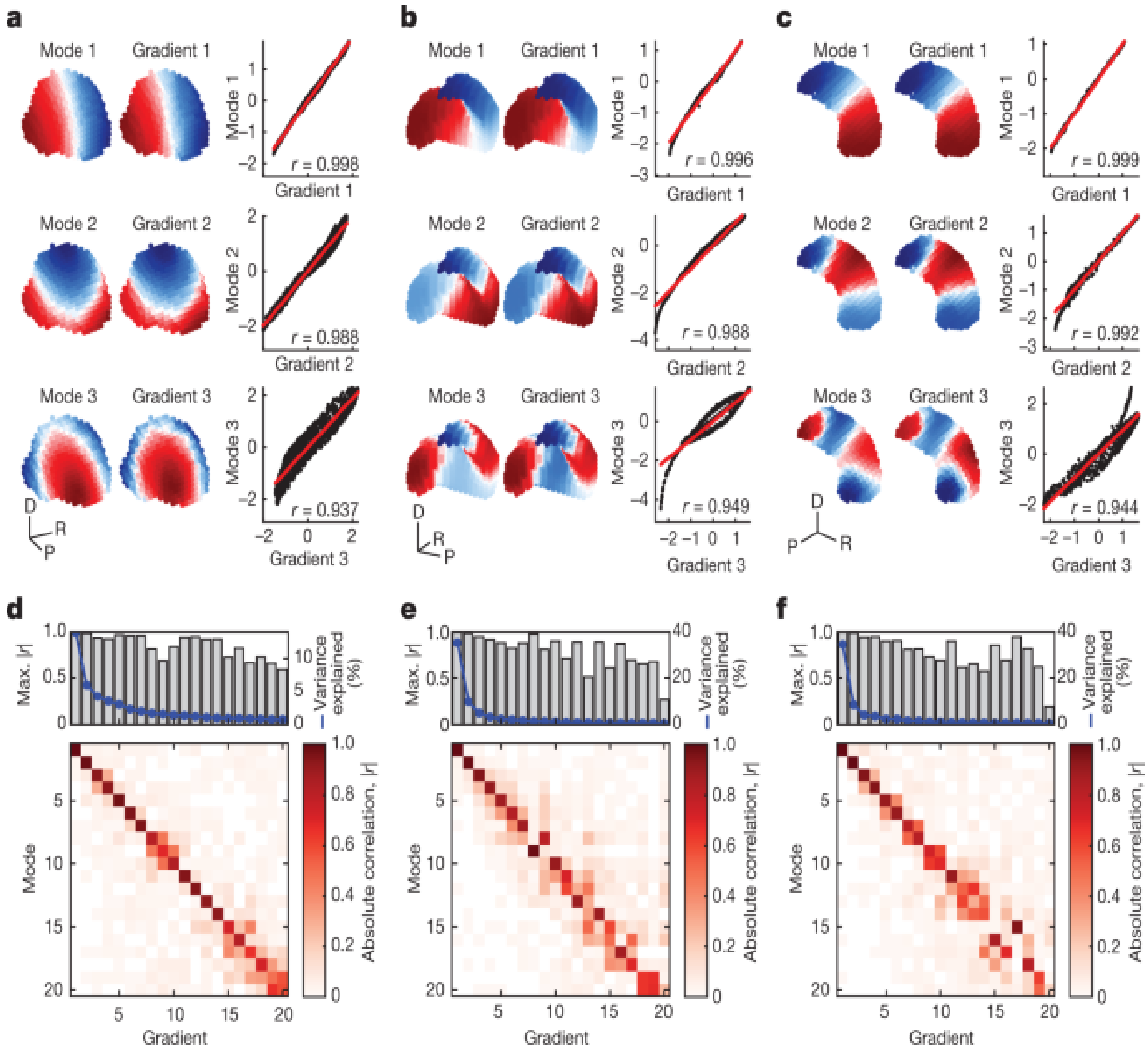
Our analyses thus far have focused on the strong coupling of geometry and dynamics in the neocortex. We next investigate this coupling in non-neocortical areas, focusing on the thalamus, striatum and hippocampus, because these structures have geometries easily captured using MRI data and their functional organization has been extensively studied[48].

We first generalize our eigenmode analysis to three-dimensional (3D) volumes (Estimating the geometric eigenmodes of non-neocortical structures in Methods), yielding geometric eigenmodes that extend spatially through the three spatial dimensions of each structure. Next, to fully capture the macroscale functional organization of these non-neocortical regions, we apply a widely used manifold learning procedure to voxel-wise FC data to obtain the key functional gradients in each structure[49] (Mapping the functional organization of non-neocortical structures in Methods). These functional gradients describe the principal axes of spatial organization

dictated by similarities in FC, thus representing the dominant modes of variation in functional organization, ordered according to the percentage of variance in FC similarity that they explain.

The spatial profiles of the first three functional gradients of the thalamus, striatum and hippocampus (accounting for 24, 50 and 47% of the variance in FC similarity, respectively) show a near-perfect match to the first three geometric eigenmodes (Fig. 5a–c; spatial correlations $r \geq 0.93$). This tight correspondence generalizes out to the first 20 modes and first 20 gradients of each structure (with the first 20 gradients respectively accounting for 49, 70 and 68% of total variance in FC similarity), with all absolute spatial correlations $|r| > 0.5$, except for the 20th gradient and 20th mode in the striatum and hippocampus (Fig. 5d–f). This strong relationship is striking given that the functional gradients are generated via a complex processing pipeline applied to fMRI-derived FC measures whereas the eigenmodes are derived simply from each structure's geometry, independent of the functional data. These findings suggest that the functional organization of non-neocortical structures derives directly from their geometric eigenmodes.

Fig. 5: Geometry shapes non-neocortical function.



a–**c**, First three geometric eigenmodes and FC-based functional gradients in the thalamus (**a**), striatum (**b**) and hippocampus (**c**). The modes and gradients are shown in 3D coordinate space with negative, zero and positive values coloured blue, white and red, respectively. D, dorsal; P, posterior; R, rightward. Scatter plots show the relationship between modes and gradients, with red lines representing linear fits with Pearson correlation coefficient *r*. **d**–**f**, Bottom, absolute correlation (|*r*|) of the first 20 geometric eigenmodes and functional gradients in the thalamus (**d**), striatum (**e**) and hippocampus (**f**). Top, highest |*r*| value obtained for each functional gradient (grey bars), taking into account order flips in geometric eigenmodes, along with the percentage of variance explained by each functional gradient (blue lines). Max., maximum.

Discussion

The dynamics of many physical systems are constrained by their geometry and can be understood as excitations of a relatively small number of structural modes[12]. Here we show that structural eigenmodes derived solely from the brain's geometry provide a more compact, accurate and parsimonious representation of its macroscale activity than alternative connectome-based models. This mode-based view of the brain further shows that both spontaneous and evoked brain activity captured by fMRI are dominated by large-scale eigenmodes with relatively long wavelengths, whose dynamics are derived from a biophysically motivated wave equation. These findings challenge the classical neuroscientific paradigm in which topologically complex patterns of interregional connectivity between discrete, specialized neuronal populations are viewed as a critical anatomical foundation for dynamics. Instead, our results indicate that a physically grounded approach that treats the brain as a continuous, spatially embedded system offers a unifying framework for understanding structural constraints on diverse aspects of macroscopic neuronal function.

The extensive comparisons of geometric eigenmodes with other anatomical (connectome and EDR eigenmodes) and statistical (PCA and Fourier) basis sets show that the superior performance of geometric eigenmodes in capturing macroscale neocortical activity is not trivially driven by generic mathematical properties of basis set expansions. Rather, this result indicates that geometry represents a fundamental anatomical constraint on dynamics. Additionally, the strong performance of the EDR eigenmodes derived from a synthetic network suggests that a homogeneous, distance-dependent connectivity with near-exponential form represents another important anatomical constraint on activity. EDR-like connectivity is mathematically embedded within the Helmholtz equation in equation (1)[6,7] (Supplementary Information 8), so the role of such connectivity is implicitly captured by the geometric eigenmodes.

The comparatively poor performance of connectome eigenmodes indicates that topologically complex connections that exist beyond a simple EDR afford minimal further benefit in obtaining eigenmodes that can accurately explain spatiotemporal patterns of cortical activity as measured with fMRI.

Our findings thus counter traditional views that emphasize intricate patterns of anatomical connections as the primary driver of coordinated dynamics[26,50,51,52]. Indeed, recent estimates indicate that long-range cortical connections are rare[53]—they may therefore represent a relatively minor perturbation of the dominant effect imposed by EDR-like connectivity. Nonetheless the topological centrality, metabolic cost and tight genetic control of such connections[54,55] suggest that they provide important functional and evolutionary advantages beyond wave-like dynamics[56] (Supplementary Information 11). The limited resolution and sensitivity to preprocessing pipelines of dMRI and fMRI data[57,58] complicate attempts to fully uncover the functional role of these connections. High-quality animal tract-tracing and electrophysiological data may be helpful in this regard.

The close coupling between geometry and dynamics is apparent in neocortical and non-neocortical structures alike, suggesting that the functional organization of regions outside the neocortex is also dominated by distance-dependent anatomical connectivity and wave dynamics, as found in recent experiments[36,59,60]. These observations indicate that geometric eigenmodes offer a simpler, more parsimonious and mechanistically informative account of putative gradients of functional organization in non-neocortical structures than the complex manifold learning procedures currently used in the literature[49]. This is because such procedures are phenomenological, providing statistical descriptions of dominant sources of variances in the data, whereas the study of structural eigenmodes derives from a generative process.

Geometric mode decomposition offers unique insights into the spatial properties of brain activation maps. Classical brain-mapping analyses focus on responses in isolated clusters of spatial locations that exceed a statistical threshold[61]. By contrast, our approach aligns with rigorously established results from physics and engineering in which perturbations of spatially continuous systems elicit system-wide responses, just as the musical notes of a violin string result from vibrations across its entire length rather than from the behaviour of a restricted segment[62]. Accordingly, the use of geometric eigenmodes indicates that, across more than 10,000 diverse maps from task-based fMRI studies, task engagement is associated predominantly with the excitation of modes with wavelengths of roughly 60 mm and longer. This

result coincides with similar observations of long-wavelength excitations in empirical EEG and evoked response potential data[30,31] and suggests that classical analyses reliant on thresholding of pointwise statistical maps obscure the spatially extended and complex patterns of activity actually evoked by a task.

Our modelling results offer insight into the physical processes that underlie the close link observed between geometry and function. In particular, the relative simplicity and superior performance of the wave model in capturing diverse aspects of spontaneous fMRI dynamics indicate that the model provides a more parsimonious account than a complex neural mass model, which treats the brain as a graph of discrete anatomical regions (nodes) coupled via the connectome (edges). This finding is consistent with experimental observations of wave dynamics in both human and animal fMRI data[63,64]. Future work could explore whether the introduction of spatial heterogeneities[65] or complex structured input[66] into the wave model further improves its accuracy in explaining diverse empirical phenomena.

Application of the wave model to mimic visual stimulation shows that waves propagating from the stimulation site segregate along the classical dorsal and ventral visual pathways, and that regional responses to the perturbation conform to a well-described hierarchy of timescales that range from rapidly responding unimodal areas to slower transmodal regions[41,44,45]. These canonical properties of hierarchical visual processing have been extensively studied for decades and are conventionally thought to be driven by complex patterns of layer-specific interregional connectivity[40,42,43], but our analysis shows that waves travelling through the cortical geometry are sufficient for the emergence of segregated, hierarchical processing streams. Thus, while our findings cannot rule out a role for complex interregional connectivity they do indicate that such connectivity is not necessary for the emergence of these macroscale dynamics.

The superior performance of geometric eigenmodes offers an immediate practical benefit because the modes can be estimated using only a mesh representation of the structure of interest, which can easily be derived using well-established, automated processing pipelines for T1w anatomical images[67]. By contrast, connectome eigenmodes require a graph-based model

of macroscopic interregional connectivity generated via complex data-processing pipelines applied to both T1w and dMRI images[58,68]; the definition of graph nodes, which is a topic of contention[69]; and the application of a thresholding procedure to remove putatively spurious connections, which our own analysis shows can affect the findings (Supplementary Figs. 6 and 7). The fact that such choices are not required to obtain the geometric eigenmodes means that they can be applied robustly and flexibly across different experimental contexts in both humans and other species[70,71], opening new avenues of research. For example, one can investigate how geometric eigenmodes vary through neurodevelopment or are disrupted in clinical disorders. Indeed, the close link we identify between geometry and function implies that interspecies differences in spatiotemporal dynamics may largely be driven by differences in brain shape. The characterization of how variations in brain geometry, both within and between species, shape brain function will be essential for understanding physical and anatomical constraints on neuronal activity.

Methods

Derivation of cortical geometric eigenmodes

If brain structure can be approximated as being constant in time, the resulting spatial and temporal dynamics can be treated separately via eigenmode decomposition[1,7], similar to the treatment of other physical systems[12]. In particular, the spatial aspect satisfies the Laplacian eigenvalue problem, which is also known as the Helmholtz equation, defined in equation (1).

For the cerebral cortex, which we consider as a two-dimensional (2D) model embedded within 3D Euclidean space, the LBO in equation (1) captures intrinsic geometry, which includes the curvature of the cortical surface[72] and is defined generally as[73,74],

$$\varDelta {\rm{:=}}\frac{1}{W}\sum _{i,j}\frac{\partial }{\partial {x}_{i}}\left({g}^{ij}W\frac{\partial }{\partial {x}_{j}}\right),$$
(2)

where x_i,x_j are the local coordinates, g^{ij} is the inverse of the inner product metric tensor ${g}_{ij}:\,=\langle \frac{\partial }{\partial {x}_{i}},\frac{\partial }{\partial {x}_{j}}\rangle$, $W := \sqrt{\det(G)}$, det denotes the determinant and $G:\,=(\,{g}_{ij})$.

We employed the LaPy python library[72,75] installed in the MASSIVE high-performance computing facility[76] to derive the geometric eigenmodes of the human cortex. Specifically, we used a triangular surface mesh representation of the midthickness human cortical surface, comprising 32,492 vertices in each hemisphere, obtained from a downsampled, left–right symmetric version of the FreeSurfer's fsaverage population-averaged template[77] (https://github.com/ThomasYeoLab/CBIG/tree/master/data/templates/surface/fs_LR_32k). This template is independent of the data sample used in all our analyses, thus obviating any concerns about circularity.

Note that the continuous LBO operates on the underlying Riemannian manifold of the surface and not directly on mesh vertices. LaPy uses the cubic finite element method on the surface mesh to achieve numerically tractable solutions of equation (1) on an interpolated smooth manifold. This distinguishes it from the discrete graph Laplacian[78], which does not encode spatial relations between points. All our analyses were focused on unihemispheric eigenmodes, but our approach can easily be extended to the whole brain because bihemispheric eigenmodes can be represented as symmetric or antisymmetric combinations of the eigenmodes derived from each hemisphere[7]; symmetric combinations correspond to mirror symmetry across the sagittal midplane and asymmetric combinations correspond to cases in which the hemispheres have the same spatial structure but with flipped signs.

The eigenvalue solutions of equation (1) are ordered sequentially according to the spatial frequency or wavelength of the spatial patterns of each eigenmode—that is, $0\le {\lambda }_{1}\le {\lambda }_{2}\le \ldots$. Note that the first eigenvalue, λ_1, is approximately equal to zero (wavelength much greater than brain size) and the corresponding eigenmode, ψ_1, is a constant function with no nodal lines (zero sets of the function). Throughout our study, we used the first 200 modes (including the constant mode, ψ_1) in

our analyses given the diminishing improvements in reconstruction accuracy observed when using an increasing number of modes (Fig. 1d).

Each eigenmode comprises spatial patterns with a specific spatial wavelength. Following ref. [7], we approximate the eigenmode wavelengths using an idealized spherical case because it is topologically comparable to the human cortex. By solving equation (1) on a sphere, degenerate solutions exist such that certain eigenmodes have the same eigenvalue and spatial wavelength—this is analogous to spherical harmonics in quantum physics. In fact, because the eigenmodes will approach the spherical harmonics in the limit of vanishing cortical folding[7], the former can be grouped together into an eigengroup with spatial wavelength[71],

$${\rm{wavelength}}=\frac{2\pi {R}_{{\rm{s}}}}{\sqrt{l(l+1)}}, \quad (3)$$

where R_s is the radius of the sphere (for the fsaverage population-averaged template used in this study, $R_s \approx 67.0$ mm) and l is the eigengroup number (the angular momentum quantum number in atomic physics). The wavelengths of the first 15 eigengroups and the eigenmodes included in the eigengroup are shown in Supplementary Table 1.

Modal decomposition of brain activity

We used the geometric eigenmodes to decompose spatiotemporal fMRI data, measured at spatial location $\mathbf{r}$ and time t, for each individual, as a weighted sum of modes,

$$y({\bf{r}},t)=\mathop{\sum }\limits_{j=1}^{N}{a}_{j}(t){\psi }_{j}({\bf{r}}), \quad (4)$$

where a_j is the amplitude of mode j in explaining the data, ψ_j is the jth mode and N is the number of modes used; we used $N = 200$ for our analyses. For spatiotemporal data—that is, recordings of spontaneous dynamics from task-free fMRI—each time frame of the data was substituted into equation (4), resulting in a time-dependent amplitude ${a}_{j}\left(t\right)$ for each

mode ψ_j. For purely spatial data—that is, task-evoked activation maps—amplitudes are independent of time such that ${a}_{j}\left(t\right)\to {a}_{j}$. In both cases, the amplitudes can be obtained by integrating over the cortical surface,

$${a}_{j}(t)=\int y({\bf{r}},t){\psi }_{j}({\bf{r}})d{\bf{r}},$$
(5)

which can be derived from equation (4) using the orthogonal property of the eigenmodes[62,79]. If there are insufficient measurements to evaluate the integral, the amplitudes can also be estimated via a statistical general linear model.

After obtaining the amplitudes, equation (4) was used to calculate the reconstructed data. We quantified the accuracy of this reconstruction by calculating the correlation between empirical and reconstructed data. For the spatiotemporal task-free data, we first parcellated the empirical and reconstructed data by taking the average of the data within discrete parcels/regions as per standard practice in the field[80], and then constructed a matrix of interregional FC by calculating the Pearson correlation coefficient of pairs of parcel time series. For task-evoked data, we applied the same parcellation on the activation maps to allow direct comparison. Finally, reconstruction accuracy for task-free data was calculated by taking the correlation of the upper triangular elements of the empirical and reconstructed FCs. For task-evoked data, reconstruction accuracy was calculated from the spatial correlation of parcellated empirical and reconstructed maps. We then took the average reconstruction accuracy across all participants.

HCP data

We used preprocessed fMRI data from HCP[27]. We did not perform any additional preprocessing steps, such as global signal removal, because the first eigenmode (considered as the global, constant mode) already explicitly captures global deviations in the data, allowing the other modes to capture functionally relevant non-global activity. We analysed data from 255 unrelated healthy individuals (aged 22–35 years, 132 females and 123

males), which is the largest HCP sample excluding twins or siblings, and with all participants having completed task-evoked and task-free resting-state data. All participants were volunteers and provided informed consent. The open-access data were acquired by the WU–Minn HCP consortium with local overseeing ethics committee approval and were shared with the authors according to HCP's data use terms. All our procedures were carried out in accordance with protocols set by these data use terms. For a detailed account of the image acquisition protocol, preprocessing pipelines and ethics oversight, see refs. [27,32].

Within the HCP dataset we analysed task-evoked fMRI measured in seven task domains (Supplementary Table 2) and task-free resting-state fMRI, which were already preprocessed by HCP. Both datasets were mapped onto the fsLR-32k CIFTI space with 32,492 vertices in each hemisphere. Further details on data and preprocessing can be found in Supplementary Information 2.1 and 2.2. We also analysed individual connectomes derived from dMRI data, as provided in ref. [81]. The connectomes represent a high-resolution weighted matrix of size 32,492 × 32,492 for each hemisphere. Further details on data and connectome construction can be found in Supplementary Information 2.3.

Cortical parcellations

The main results of the study analysed task-evoked activation maps and task-free FC data parcellated into discrete regions. We presented results using the HCP–MMP1 parcellation with 180 regions per hemisphere (we term this the Glasser360 parcellation), which reflects sharp areal boundaries based on the combination of cortical architecture, function, connectivity and topography[28]. The parcellation was on the fsLR-32k space as provided by HCP. To test the robustness of our results (Supplementary Fig. 2 and Extended Data Figs. 2 and 3) we also performed our analysis using parcellations provided by Schaefer et al.[82] on the fsLR-32k space at varying resolution (100, 200, 400, 600, 800 and 1,000 parcels across both hemispheres; we refer to these as, respectively, Schaefer100, Schaefer200, Schaefer400, Schaefer600, Schaefer800 and Schaefer1000 parcellations).

Derivation of connectome eigenmodes

Connectome eigenmodes were derived according to previous methods[29] to enable comparison with previous findings. Note that, in previous studies[19,20,29,83], connectome eigenmodes have been referred to as connectome harmonics but we use the term eigenmodes here because the term harmonics implies integer frequency ratios, which is not necessarily guaranteed for brain-derived modes.

We obtained high-resolution maps of connectivity measured with dMRI tractography as described in ref. [81], in which the connectivity of each of the 32,492 vertices in the cortical surface mesh was estimated by tracing streamlines from each point until they terminated at some other point. Connection weights between vertices (considered as nodes) were estimated as the number of interconnecting streamlines without the need for normalization[84]. Tractography was performed on individuals from HCP (see Supplementary Information 2.3 for further details on the data and tractography method). From the tractography data we combined the individual weighted connectivity matrices of size 32,492 × 32,492 to generate a group-averaged connectome, $W_{\rm connectome}$, with weights representing the average number of streamlines. We then generated a binary adjacency matrix, $A_{\rm local}$, that captures the discrete representation of local spatial relations between points in the cortical surface mesh model constructed by connecting two vertices that are direct neighbours in the mesh. These links are intended to capture local, very-short-range connectivity that cannot be resolved by traditional dMRI tractography[29].

Following ref. [29], the group-averaged weighted connectome, $W_{\rm connectome}$, was thresholded to remove the smallest weights such that the number of connections was fourfold greater than $A_{\rm local}$. The resulting thresholded matrix was binarized to obtain the group adjacency matrix, $A_{\rm connectome}$. Finally we generated a merged adjacency matrix ${A}_{{\rm{C}}}={A}_{{\rm{local}}}| | {A}_{{\rm{connectome}}}$ (with matrix of size 32,492 × 32,492; || is the logical OR operator), which captures both local vertex-to-vertex connectivity and complex short- and long-range connections as measured empirically. Note that the connection density of the adjacency matrix, $A_{\rm C}$, that resulted from the thresholding process described above was 0.10%.

The connectome eigenmodes were obtained by solving the eigenvalue problem,

$${L}^{{\prime} }\psi =-\lambda \psi ,$$
(6)

where L' is the normalized graph Laplacian, a discrete counterpart of the LBO. The normalized graph Laplacian is related to the unnormalized graph Laplacian, L, as $({L}^{{\prime} }={D}^{-1/2}L{D}^{-1/2})$, with L defined according to previous work[34],

$$L=\frac{1}{2}\left[\left(D-{A}_{{\rm{C}}}\right)+{(D-{A}_{{\rm{C}}})}^{T}\right],$$
(7)

where D is the diagonal degree matrix and superscript T denotes matrix transpose. As with geometric eigenmodes, the eigenvalue solutions of equation (6) form the sequence $0\le {\lambda }_{1}\le {\lambda }_{2}\le \ldots$. Note also that the use of a high-resolution, vertex-level connectome results in connectome eigenmodes spanning a space with dimensions (number of modes) equal to the number of vertices, allowing fair comparison with geometric eigenmodes.

Derivation of EDR eigenmodes

Exponential distance rule eigenmodes were derived also by solving equation (6). However, in this case the non-normalized graph Laplacian, L, was defined using a synthetically constructed 32,492 × 32,492 adjacency matrix, A_E, that follows a stochastic EDR. To construct A_E we used the group-averaged, unthresholded weighted connectome, $W_{\text{connectome}}$, from the previous section. We then fitted the variation of weights as a function of Euclidean distance, d, between vertices in the cortical surface by an exponential function of the form ${e}^{-\alpha d}$, where α is a scale exponent parameter. The fitting was performed in MATLAB 2019b using a nonlinear least-squares method, resulting in an optimal empirical parameter value of ${\alpha }_{{\rm{empirical}}}=0.12$, consistent with previous estimates based on the connection probability versus distance function[85]. We

then generated a random, binary adjacency matrix following a stochastic EDR wiring process in which two vertices were connected with probability ${e}^{-\alpha d}$ with $\alpha ={\alpha }_{{\rm{empirical}}}=0.12$.

The adjacency matrix, A_E, generated by instances of the EDR model had a connection density of 1.55%, which was substantially higher than the 0.10% density of A_{C}. We therefore constructed another version of connectome eigenmodes that thresholded the group-averaged empirical connectome to achieve a final A_{C} with a density that matched A_E to allow fair comparison (we termed this density-matched connectome eigenmodes). Supplementary Figs. 6 and 7 present a more thorough evaluation of the performance of connectome eigenmodes as a function of network connection density.

Comparisons with statistical basis sets

The geometric, connectome and EDR eigenmode basis sets were all derived from a generative model that accounts for how brain function emerges from anatomy. This approach contrasts with the statistical basis sets commonly used in the literature, which can efficiently summarize the data but offer no insights into the underlying generative process. We evaluated the performance of geometric eigenmodes with respect to two statistical basis sets, one derived from PCA of the functional data themselves and the other based on a Fourier spatial basis set. Further details of these analyses are provided in Supplementary Information 5 and 6, respectively.

Modal power spectra of task-evoked activation maps

To investigate the spectral content of task-evoked activation maps we calculated their modal power spectra using the modal decomposition in equation (4) and taking the absolute square of the amplitudes, *a*. This is analogous to calculation of the temporal power spectral density from Fourier analysis. We then normalized the power in mode *j* with respect to the total power in all modes, such that

$${P}_{j}=\frac{{|{a}_{j}|}^{2}}{{\sum }_{j=1}^{N}{|{a}_{j}|}^{2}}.$$
(8)

We calculated the modal power spectra of two sets of task-evoked data. The first set comprises unthresholded activation maps from the HCP task-evoked data (Supplementary Information 2.1). Before perfoming spectral analysis we took the group average of the activation maps across 255 individuals and then analysed the power spectrum of the group-averaged activation map of each task contrast. The results in Fig. 3a and Extended Data Fig. 8a show the mean power spectrum across the 47 HCP task-contrast maps. Contrast-specific power spectra for the seven key HCP task contrasts are shown in Extended Data Fig. 7.

The second set comprises 10,000 activation maps from 1,178 independent experiments from the NeuroVault repository (https://neurovault.org/)[33]. We used the python module Nilearn[86] to retrieve activation maps from NeuroVault that were unthresholded and with a modality tag of fMRI-BOLD. We projected the activation maps from volume onto the fsLR-32k CIFTI space to match the HCP data using Nilearn (via the function nilearn.surface.vol_to_surf). We then analysed the power spectrum of each activation map. The results in Fig. 3a and Extended Data Fig. 8a show the mean power spectrum across the 10,000 NeuroVault maps.

We then compared the power spectra of the empirical maps with those of surrogate random maps with varying levels of smoothing to further investigate the relevance of long-wavelength modes. In particular, we generated 10,000 random maps in volume space, which we smoothed at kernel sizes with full-width at half-maximum ranging from 0 to 50 mm, and projected onto the fsLR-32k CIFTI space. We then calculated the mean square logarithmic error (MSLE) of the power spectra between the empirical and surrogate maps. Further details of this analysis and the MSLE measurements are provided in Supplementary Information 7.

Contributions of long- and short-wavelength modes

To understand the contributions of long- and short-wavelength geometric eigenmodes in the reconstruction of task-evoked activation maps, we sequentially removed modes before performing the reconstruction process. Specifically, we started by reconstructing the seven key HCP contrast maps using 200 geometric eigenmodes and calculated the reconstruction accuracy (that is, correlation between empirical and reconstructed maps), serving as our baseline. We then performed an incremental, sequential removal of modes starting from long-wavelength modes (that is, removal of mode 1, modes 1–2, modes 1–3, modes 1–4, …, modes 1–200) and calculated reconstruction accuracy at every increment. We repeated the same procedure but starting from short-wavelength modes (that is, removal of mode 200, modes 199–200, modes 198–200, modes 197–200, …, modes 1–200).

NFT wave model

As stated above, an implication of the superior performance of geometric eigenmodes is that neuronal activity is dominated by wave dynamics as predicted by NFT. To investigate whether wave dynamics can explain complex spatiotemporal patterns of neuronal activity, we implemented a simple NFT wave model in which dynamics are described by an isotropic damped wave equation without regeneration[5,6,87],

$$\left[\frac{1}{{\gamma }_{s}^{2}}\frac{{{\rm{\partial }}}^{2}}{{\rm{\partial }}{t}^{2}}+\frac{2}{{\gamma }_{s}}\frac{{\rm{\partial }}}{{\rm{\partial }}t}+1-{r}_{s}^{2}{{\rm{\nabla }}}^{2}\right]\varphi ({\bf{r}},t)=Q({\bf{r}},t),$$

(9)

where $\varphi ({\bf{r}},t)$ is the neural activity at location **r** and time t, Q is an external input, γ_s is the damping rate and r_s is the spatial length scale of the wave propagation (conceptually related to the α-length scale of the stochastic EDR wiring process in the derivation of EDR eigenmodes). This form tells us that an impulse input will produce an activity that dissipates at a rate of γ_s and propagates at a velocity of $\gamma_s r_s$. Here we treated γ_s as a fixed parameter with the value of 116 s^{-1} taken from electrophysiological estimates[21] and r_s as a free parameter. We applied the model on the cortical midthickness surface mesh with 32,492 vertices per hemisphere to solve the activity at each vertex. Note that the propagation of activity between points is governed by their white-matter connectivity, with strength that decays approximately exponentially with distance (Supplementary Fig. 10). This distance dependence is more apparent when equation (9) is converted into its equivalent integral form (see Supplementary Information 8 for details).

Neural mass model

We compared the performance of our simple NFT wave model with a biophysical large-scale neural mass model in which mesoscopic dynamics emerge from the interactions of neural populations (that is, neural masses) coupled via an empirical anatomical connectivity[18]. In a typical neural mass

model, each brain region i has its own mean-field population dynamics and its temporal activity, S_i, is defined by the general equation,

$$\frac{{\rm{d}}{S}_{i}}{{\rm{d}}t}=f({\bf{S}},{\theta }_{i},C,G),$$
(10)

where f is a function describing the evolution of the region's activity. The function depends on the activity of other regions **S**, local population parameters θ_i, anatomical connectivity between regions C (a parcellated version of the vertex-resolution connectome, $W_{\text{connectome}}$, used to derive the connectome eigenmodes) and global coupling parameter G that scales the connectivity between regions.

There are several whole-brain neural mass models available in the literature that we can use[88], from simple phase oscillator models (for example, the Kuramoto model[89]) to more complex biophysical population models (for example, the Wilson–Cowan model[90]). All these models follow the form of equation (10), especially their reliance on an anatomical interregional connectivity matrix, C. Here we focus on one widely used neural mass model, the BEI model[38,65,91,92]. The BEI model uses a mean-field approach to approximate local population dynamics in each brain region, which are coupled via an anatomical connectivity matrix derived from dMRI. Each brain region i comprises interacting populations of excitatory (E) and inhibitory (I) neurons governed by the following nonlinear stochastic differential equations,

$$\frac{{\rm{d}}{S}_{i}^{(E)}}{{\rm{d}}t}=\,-\,\frac{{S}_{i}^{(E)}}{{\tau }_{E}}+\left(1-{S}_{i}^{(E)}\right)\gamma {r}_{i}^{(E)}+\sigma {\nu }_{i}(t),$$
(11)
$$\frac{{\rm{d}}{S}_{i}^{(I)}}{{\rm{d}}t}=\,-\,\frac{{S}_{i}^{(I)}}{{\tau }_{I}}+{r}_{i}^{(I)}+\sigma {\nu }_{i}(t),$$
(12)
$${r}_{i}^{(E)}={H}^{(E)}\left({I}_{i}^{(E)}\right)=\frac{{a}_{E}{I}_{i}^{(E)}-{b}_{E}}{1-\exp \left[-{d}_{E}\left({a}_{E}{I}_{i}^{(E)}-{b}_{E}\right)\right]},$$
(13)

$${r}_{i}^{(I)}={H}^{(I)}\left({I}_{i}^{(I)}\right)=\frac{{a}_{I}{I}_{i}^{(I)}-{b}_{I}}{1-\exp \left[-{d}_{I}\left({a}_{I}{I}_{i}^{(I)}-{b}_{I}\right)\right]},$$
(14)
$${I}_{i}^{(E)}\left(t\right)={I}^{{\rm{ext}}}+{W}_{E}{I}_{0}+{w}_{EE}{S}_{i}^{(E)}\left(t\right)+GJ\sum _{j}{C}_{ij}{S}_{j}^{(E)}\left(t\right)-{w}_{IE}{S}_{i}^{(I)}\left(t\right),$$
(15)
$${I}_{i}^{(I)}\left(t\right)={W}_{I}{I}_{0}+{w}_{EI}{S}_{i}^{(E)}\left(t\right)-{S}_{i}^{(I)}\left(t\right),$$
(16)

where ${S}_{i}^{(E,I)}$, ${r}_{i}^{(E,I)}$ and ${I}_{i}^{\left(E,I\right)}$ represent the synaptic gating variable, firing rate and total input current, respectively, for *E* and *I* populations. The parameters $\tau_{E,I}$ are the time constants, γ is a kinetic rate constant and $v_i(t)$ is a time-varying random Gaussian input with standard deviation σ. The functions $H^{(E,I)}$ are sigmoidal neuronal response functions transforming total input currents ${I}_{i}^{(E\,,I)}$ into firing rates ${r}_{i}^{(E,I)}$, which are parametrized by the gain factors $a_{E,I}$, threshold currents $b_{E,I}$ and curvature parameters $d_{E,I}$. In equations (15) and (16), ${I}_{i}^{{\rm{ext}}}$ is the external input, I_0 is the local input current scaled by W_E and W_I for the excitatory and inhibitory populations, respectively, w_{EE} is the excitatory–excitatory strength, w_{EI} is the excitatory–inhibitory strength, w_{IE} is the inhibitory–excitatory strength, *G* is the global coupling parameter, *J* is the effective *N*-methyl-d-aspartate (NMDA) conductance and C_{ij} is the structural connectivity strength between regions *i* and *j* estimated from dMRI.

Overall, the BEI model has 15 fixed parameters and four free parameters, as detailed in Supplementary Table 3. The values of the fixed parameters were taken from ref. [91]. To allow direct comparison we also used the discretized connectome data provided by ref. [91] to define the interregional structural connectivity matrix, *C*. These connectome data were derived from minimally preprocessed dMRI data of 334 unrelated HCP subjects and constructed via the probabilistic tractography tool of FMRIB Software

Library (FSL)[93]. We refer readers to ref. [91] for further details of data processing and connectome construction. Our results did not change when the parcellated version of our connectome data was used. We note that numerical solutions of NFT equations (such as in equation (9)) also spatially discretize the cortex, but into a very fine array of points, before integrating whatever temporal dynamics have been chosen at each point with connectivity that is correct in the continuum limit. As such, neural mass models are approximations to the more general NFT approach[94,95]. We also reiterate that local dynamics do not affect the spatial eigenfunctions of the NFT wave model.

Haemodynamic model

To simulate fMRI data we transformed the neural activity generated by the NFT wave and BEI neural mass models to a blood oxygen-level dependent (BOLD) signal using the well-established Balloon–Windkessel haemodynamic model[96]. Note that, although this model is a simple approximation to more detailed models of the physiological haemodynamic processes underlying the BOLD signal[97,98,99], we use this approximation here to allow direct comparison with the vast majority of modelling studies in the literature[65,91,92,100,101]. The BOLD–fMRI signal in each vertex or brain region *i* is governed by the following differential equations,

$$\frac{{\rm{d}}{z}_{i}}{{\rm{d}}t}={N}_{i}(t)-\kappa {z}_{i}(t)-\gamma [\,{f}_{i}(t)-1],$$
(17)
$$\frac{{\rm{d}}{f}_{i}}{{\rm{d}}t}={z}_{i}(t)\,,$$
(18)
$$\frac{{\rm{d}}{v}_{i}}{{\rm{d}}t}=\frac{1}{\tau }[\,{f}_{i}(t)-{v}_{i}^{1/\alpha }(t)],$$
(19)
$$\frac{{\rm{d}}{q}_{i}}{{\rm{d}}t}=\frac{1}{\tau }\left\{\frac{{f}_{i}(t)}{\rho }[1-{(1-\rho)}^{1/{f}_{i}(t)}]-{v}_{i}^{1/\alpha -1}(t)\right\},$$
(20)
$$\frac{{\rm{d}}{y}_{i}}{{\rm{d}}t}={V}_{0}\left\{{k}_{1}\left[1-{q}_{i}(t)\right]+{k}_{2}\left[1-\frac{{q}_{i}(t)}{{v}_{i}(t)}\right]+$$

{k}_{3}\left[1-{v}_{i}(t)\right]\right\},$$
(21)

where z_i, f_i, v_i, q_i and y_i are the vasodilatory signal, blood inflow, blood volume, deoxyhaemoglobin content and BOLD signal variables, respectively. The variable N_i represents the neural activity generated by the wave and BEI neural mass models. For the wave model, $N(t)$ is $\phi(t)$, and for the BEI neural mass model, $N_i(t)$ is $S_i^{(E)}(t)$. The model parameters and their values taken from previous work[96,102] were as follows: $\kappa = 0.65\ \mathrm{s}^{-1}$ is the signal decay rate, $\gamma = 0.41\ \mathrm{s}^{-1}$ is the flow-dependent elimination rate, $\tau = 0.98$ s is the haemodynamic transit time, $\alpha = 0.32$ is the Grubb's exponent, $\rho = 0.34$ is the resting oxygen extraction fraction, $V_0 = 0.02$ is the resting blood volume fraction and $k_1 = 3.72$, $k_2 = 0.53$ and $k_3 = 0.53$ are 3T fMRI parameters.

Modelling resting-state dynamics

We used the NFT wave and BEI models, together with the haemodynamic model, to estimate various spontaneous FC properties. For the wave model we solved equation (9) with a white noise input to mimic the absence of any structured stimulus, following previous studies[21,103]. We combined the resulting solution with the haemodynamic model in equations (17–21) to simulate the BOLD–fMRI signal. The simulated BOLD signal was downsampled to a sampling interval of 0.72 s with 1,200 time frames to match the resting-state HCP data described in Supplementary Information 2.2. Finally we parcellated the simulated BOLD signal using the HCP–MMP1 parcellation and calculated the correlation matrix representing the model FC.

For the BEI model we can use the method described above to calculate the model FC starting from the numerical solutions of equation (11) for each brain region. However, it has been shown that one can linearize the equations of the BEI and haemodynamic models to obtain an analytic approximation of the model FC[91]. Due to the large number of model parameters in the BEI model we used this analytic approximation in this

study because it allows more comprehensive and computationally efficient model fitting. See ref. [91] for further details.

We fitted the model FCs to empirical FCs (also parcellated using the HCP–MMP1 parcellation) of the same HCP participants used in all our analyses by fitting each model's free parameters. We divided the participant sample into training and test sets, each with 125 individuals, to enable out-of-sample evaluation of model performance and to avoid overfitting. In particular, we fitted the model parameters on the training set data and used the fitted model parameters to predict the test set data. Model fitting and performance evaluation were based on three widely used FC-related metrics[65,92]: edge FC, node FC and FCD.

The edge FC metric was calculated by taking the Pearson correlation of the upper triangular elements (that is, the strength of FC edges) of the *z*-transformed model and empirical FCs. We only took the upper triangular elements because the FC values are symmetric with respect to the diagonal. Higher correlations represent a better fit between model and data.

The node FC metric was calculated by taking the Pearson correlation of the average FC strength of each brain region in the model and empirical FCs. The average FC strength of a brain region *i* was defined as $\frac{1}{n}{\sum }_{j=1}^{n}{{\rm{F}}{\rm{C}}}_{ij}$, where *n* is the total number of brain regions. Once again, higher correlations indicate better model fits.

The FCD metric captures the spatiotemporal statistics of resting-state activity and was calculated as follows[92]. The time series at each region *i* was filtered between 0.04 and 0.07 Hz using a second-order Butterworth filter; this band was based on ref. [104] and its inclusion was motivated by its functional relevance to the brain[105,106]. The time series was then Hilbert (*H*) transformed to calculate the quantity ${y}_{i}\left(t\right)={x}_{i}\left(t\right)+j{H}_{i}(t)$, where *j* is the imaginary number. The instantaneous complex argument ${\theta }_{i}\left(t\right)={\tan }^{-1}\left[{H}_{i}(t)/{x}_{i}(t)\right]$ was then computed. The level of synchrony between regions *i* and *j* at time *t*, $\Delta (i,j,t)$, was calculated as

$$\Delta \left(i,j,t\right)=\cos \left[{\theta }_{i}\left(t\right)-{\theta }_{j}(t)\right]. \tag{22}$$

Note that we do not term ${\theta }_{i}\left(t\right)$ a phase, as is sometimes done in the literature, because our signals are broadband whereas this interpretation is applicable only to signals that are nearly monochromatic. We computed this quantity for comparison with previous work[65,92], then we calculated the similarity of global synchrony, φ_{uv}, between two time instances, τ_u and τ_v, with φ_{uv} defined as

$${\phi }_{uv}=\frac{1}{{d}_{u}{d}_{v}}\sum _{i > j}\Delta (i,j,{\tau }_{u})\Delta (i,j,{\tau }_{v}), \tag{23}$$

where

$${d}_{x}=\sqrt{\sum _{i > j}{[\Delta (i,j,{\tau }_{x})]}^{2}}. \tag{24}$$

Here, φ_{uv} is the FCD, which is a symmetric time × time matrix interpreted as the similarity of global synchrony between times τ_u and τ_v. We then compared the distributions of the upper triangular elements of the model and empirical FCD estimates, concatenated across all individuals or model realizations as per ref. [92], to better capture the general fluctuations of the data. The FCD distributions of the model and the data were compared using the KS statistic, with lower KS statistic indicating better model fit.

The wave model has one free parameter—the spatial length scale of wave propagation, r_s—that was optimized to fit the empirical fMRI data. Specifically, we constructed a vector of 20 values of r_s evenly spaced between 10 and 100 mm. For each r_s value the FC metrics described above were calculated. The optimization landscapes are shown in Extended Data Fig. 10. We took the value of r_s that minimized the FCD KS statistic because this metric has been found to be the most stringent benchmark for model–data comparisons among the three metrics discussed above[92]. This

procedure resulted in an optimized value of $r_s = 28.9$ mm, which is smaller than those obtained previously in analyses of EEG data[21], possibly due to either the haemodynamic processes limiting neural activity propagation to neighbouring regions or our focus here on cortico-cortical dynamics[97,98].

The BEI neural mass model has 15 fixed parameters and four free parameters—that is, w_{EE}, w_{EI}, w_{IE} and G—that were optimized to fit the data, following ref. [91]. This procedure resulted in optimized parameters $w_{EE} = 9.80$, $w_{EI} = 1.48$, $w_{IE} = 7.13$ and $G = 6.87$. Further details of the optimization process can be found in Supplementary Information 9.

Measurement of time-lagged properties of resting-state dynamics

In addition to the FC-based metrics used for model fitting and evaluation of model performance, we investigated whether the wave and neural mass models could also capture the temporal properties of propagated activity. In particular we analysed the lag structure (or lag threads) of resting-state BOLD–fMRI time courses, as proposed by refs. [39,107]. We briefly discuss the algorithm for calculation of the lag structure of both empirical and simulated fMRI data in Supplementary Information 10 and refer readers to previous articles[39,107] for further details. Note that there are several other ways of characterizing the spatiotemporal properties of resting-state activity[37,108,109], but the method currently chosen is sufficient for our purposes.

Modelling stimulus-evoked dynamics

We also evaluated the degree to which the wave model could capture classical properties of evoked neural responses. Specifically we used the optimized wave model (with $r_s = 28.9$ mm) to investigate neural dynamics in response to a stimulus applied to the left V1. In particular, the stimulus $Q(\mathbf{r},t)$ was a 1 ms pulse ($t = 1$–2 ms) with magnitude 20 s^{-1} (the results are robust to changes in amplitude) restricted to vertices in the cortical surface that fall within the V1 region as defined by the HCP–MMP1 parcellation.

We performed the simulation over a 100 ms time period with 0.1 ms resolution. Finally we parcellated the activity $\varphi ({\bf{r}},t)$ into 180 regions using the HCP–MMP1 parcellation.

We compared the activity profile (amplitude versus time) of each brain region, focusing on the time for the activity to reach peak amplitude (that is, time to peak). Specifically we investigated whether the temporal precedence of activity follows the human visual cortical hierarchy from visual to frontal cortices, which includes the following 17 brain regions: V1, V4, 7m, 7Am, TE1p, 7AL, 24dd, 2, 24dv, 8BM, 10r, 10v, 8BL, 10pp, 10d, 9-46v and 9-46d. These regions closely resemble areas previously identified in the macaque neocortex visual hierarchy using tract-tracing data and nonlinear network modelling[41].

We also compared the regional estimates of peak response times to T1w:T2w values, which is a non-invasive measure sensitive to intracortical myelin content[46] and a good proxy for cortical hierarchy rank[110]. The myelin map in fsLR-32k space was obtained from the HCP dataset[27,28] and then parcellated using the HCP–MMP1 parcellation. We quantified the relationship between the regional values of time to peak and myelin content via Spearman rank correlation. The statistical significance of the correlation was assessed by comparison with a null distribution of 10,000 correlation values obtained using a spatially constrained spin-test approach[111,112]. This approach calculates the correlation between one map and random spatially rotated versions of the other map, thereby preserving the spatial relationship of the parcels in the map. The resulting P value, $P_{\rm spin}$, is the fraction of null correlation values greater than the empirical correlation value. Finally we repeated the above statistical test including all brain regions (that is, not restricted to visual hierarchy brain regions) to ensure that our findings were not driven by our particular selection of regions of interest (ROIs). This is a more conservative test because not all brain regions show strong evoked responses to visual stimulation.

Estimation of the geometric eigenmodes of non-neocortical structures

We extended our eigenmode analysis to regions outside the neocortex, focusing in particular on the subcortex (thalamus and striatum) and archicortex (hippocampus). Unlike the cortical ribbon, which can be modelled as a 2D sheet, these structures are solid 3D objects. We therefore calculated the geometric eigenmodes using a tetrahedral mesh rather than the surface-based triangular mesh to account for the full 3D geometry of the non-neocortical structures[72], as outlined below.

We first used the probabilistic Harvard–Oxford subcortical atlas (https://fsl.fmrib.ox.ac.uk/fsl/fslwiki/Atlases) to generate a volumetric binary mask in each hemisphere for the thalamus, striatum and hippocampus. Voxels with a probability of 25% or more of belonging to these structures were included in the masks. We used FreeSurfer's mri_mc function, which implements a marching-cubes algorithm, to construct a 2D surface by tessellating the volumetric masks, followed by the Gmsh software (https://gmsh.info/) to convert the 2D surface into a 3D tetrahedral mesh. We then used the LaPy python library to solve equation (1) on the tetrahedral mesh of each non-neocortical structure to obtain the eigenmodes and their corresponding eigenvalues. Finally we projected the eigenmodes in tetrahedral space back into the natural volumetric space via interpolation. Hence, the resulting eigenmodes spatially vary through the 3D voxels comprising each non-neocortical structure's volume. For this part of the study we discarded the first constant mode and used the next 20 modes in our analyses.

Mapping the functional organization of non-neocortical structures

The signal-to-noise ratio of fMRI is generally weaker in non-neocortical structures than in the neocortex[113]. Moreover, fine-grained task activations are often hard to resolve in these smaller structures due to the limited resolution of fMRI and the spatial smoothing induced by common fMRI processing methods. Thus, to efficiently map the functional organization of the thalamus, striatum and hippocampus we used connectopic mapping[49] of the resting-state fMRI signals in each structure to obtain their dominant functional modes (often called gradients, although the patterns capture spatial variations in pointwise similarities of FC profiles). This technique,

and related procedures, have been extensively used in past work to study the functional organization of these structures[114,115,116].

We applied connectopic mapping to the volumetric voxel-wise resting-state fMRI data of the HCP individuals following the procedure described in ref. [49]. Specifically, for each individual and non-neocortical ROI (that is, thalamus, striatum and hippocampus) we constructed an ROI time-series data matrix, A, of size $T \times N$, where T is the number of time frames and N is the number of voxels in the ROI. Similarly, for each individual we constructed the grey matter time-series data matrix, B, of size $T \times M$, where M is the number of grey matter voxels outside the ROI. Because M is generally large (10^3 or above), we reduced the dimensionality of B using singular value decomposition to construct the matrix $\widetilde{B}$ of size $T \times (T-1)$. The connectivity fingerprint of every voxel within the ROI was then calculated using the Pearson correlation of A and $\widetilde{B}$ to obtain the matrix, C, of size $N \times (T-1)$. Similarity in the connectivity fingerprints between each pair of ROI voxels was calculated using the η^2 coefficient[117], resulting in the matrix, S, of size $N \times N$. The η^2 coefficient represents the fraction of variance in one connectivity profile accounted for by the variance in another. We then took the average S across all individuals. A nonlinear manifold learning procedure using the Laplacian Eigenmaps algorithm[118] was applied to the S matrix of each ROI to calculate its eigenvectors and eigenvalues. The eigenvectors represent the functional patterns, termed functional gradients, of the ROI, ordered according to the variance in FC similarity they explain. We only analysed the first 20 non-constant gradients to enable direct comparisons with the geometric eigenmodes. Typical applications of FC-based gradient analysis rarely consider more than the first five gradients[49,119].

We analysed the correspondence between the geometric eigenmodes and functional gradients in each non-neocortical structure by taking their absolute spatial correlations, given that the signs of the modes and gradients are arbitrary. Moreover, the ordering of the geometric eigenmodes within the same eigengroup can change (Supplementary Table 1; for example, the first eigengroup can have order flips among modes 2–4) and hence one-to-one correspondence between the indices of geometric eigenmodes and functional gradients is not guaranteed. We therefore examined all possible

pairwise correlations between the modes and gradients and evaluated correspondence with respect to the mode that maximally correlated with each gradient. The maximal order differences observed between geometric modes and functional gradients were one, eight and five out of 20 for the thalamus, striatum and hippocampus, respectively.

Reporting summary

Further information on research design is available in the Nature Portfolio Reporting Summary linked to this article.

Data availability

Raw and preprocessed HCP data can be accessed at https://db.humanconnectome.org/. NeuroVault data can be accessed at https://neurovault.org/. Source data to replicate the results of the study are openly available at https://github.com/NSBLab/BrainEigenmodes and https://osf.io/xczmp/.

Code availability

Computer codes used to calculate the eigenmodes, analyse results and reproduce the figures of the study are openly available at https://github.com/NSBLab/BrainEigenmodes.

References

1. Nunez, P. L. *Neocortical Dynamics and Human EEG Rhythms* (Oxford Univ. Press, 1995).

2. Bullmore, E. & Sporns, O. Complex brain networks: graph theoretical analysis of structural and functional systems. *Nat. Rev. Neurosci.* **10**, 186–198 (2009).

3. Yuste, R. From the neuron doctrine to neural networks. *Nat. Rev. Neurosci.* **16**, 487–497 (2015).

4. Wright, J. J. & Liley, D. T. J. Simulation of electrocortical waves. *Biol. Cybern.* **72**, 347–356 (1995).

5. Jirsa, V. & Haken, H. Field theory of electromagnetic brain activity. *Phys. Rev. Lett.* **77**, 960–963 (1996).

6. Robinson, P. A., Rennie, C. J. & Wright, J. J. Propagation and stability of waves of electrical activity in the cerebral cortex. *Phys. Rev. E* **56**, 826–840 (1997).

7. Robinson, P. A. et al. Eigenmodes of brain activity: neural field theory predictions and comparison with experiment. *NeuroImage* **142**, 79–98 (2016).

8. Gabay, N. C. & Robinson, P. A. Cortical geometry as a determinant of brain activity eigenmodes: neural field analysis. *Phys. Rev. E* **96**, 032413 (2017).

9. Yang, Y., Wang, W., Lou, Y., Yin, J. & Gong, X. Geometric and amino acid type determinants for protein-protein interaction interfaces. *Quant. Biol.* **6**, 163–174 (2018).

10. Braitenberg, V. & Schüz, A. *Cortex: Statistics and Geometry of Neuronal Connectivity* (Springer-Verlag, 1998).

11. Damoiseaux, J. S. & Greicius, M. D. Greater than the sum of its parts: a review of studies combining structural connectivity and resting-state functional connectivity. *Brain Struct. Funct.* **213**, 525–533 (2009).

12. Melrose, D. B. & McPhedran, R. C. *Electromagnetic Processes in Dispersive Media* (Cambridge Univ. Press, 1991).

13. Nozari, E. et al. Is the brain macroscopically linear? A system identification of resting state dynamics. Preprint at https://arxiv.org/abs/2012.12351 (2020).

14. Jones, E. G. Golgi, Cajal and the neuron doctrine. *J. Hist. Neurosci.* **8**, 170–178 (1999).

15. Brodmann, K. Beiträge zur histologischen lokalisation der grosshirnrinde, VI: mitteilung: die cortexgliederung des menschen [German]. *J. Psychol. Neurol.* **10**, 231–246 (1908).

16. Broca, P. Remarques sur le siège de la faculté du langage articulé, suivies d'une observation d'aphémie (perte de la parole) [French]. *Bull. Mém. Soc. Anat. Paris* **6**, 330–357 (1861).

17. Fornito, A., Zalesky, A. & Bullmore, E. T. *Fundamentals of Brain Network Analysis* (Academic Press, 2016).

18. Breakspear, M. Dynamic models of large-scale brain activity. *Nat. Neurosci.* **20**, 340–352 (2017).

19. Atasoy, S., Donnelly, I. & Pearson, J. Human brain networks function in connectome-specific harmonic waves. *Nat. Commun.* **7**, 10340 (2016).

20. Preti, M. G. & Van De Ville, D. Decoupling of brain function from structure reveals regional behavioral specialization in humans. *Nat. Commun.* **10**, 4747 (2019).

21. Robinson, P. A., Rennie, C. J., Rowe, D. L., O'Connor, S. C. & Gordon, E. Multiscale brain modelling. *Philos. Trans. R. Soc. B Biol. Sci.* **360**, 1043–1050 (2005).

22. Robinson, P. A. Physical brain connectomics. *Phys. Rev. E* **99**, 012421 (2019).

23. Wang, X. J. & Kennedy, H. Brain structure and dynamics across scales: in search of rules. *Curr. Opin. Neurobiol.* **37**, 92–98 (2016).

24. Roberts, J. A., Perry, A., Roberts, G., Mitchell, P. B. & Breakspear, M. Consistency-based thresholding of the human connectome. *NeuroImage* **145**, 118–129 (2017).

25. Gabay, N. C., Babaie-Janvier, T. & Robinson, P. A. Dynamics of cortical activity eigenmodes including standing, traveling, and rotating

waves. *Phys. Rev. E* **98**, 042413 (2018).

26. Honey, C. J., Kötter, R., Breakspear, M. & Sporns, O. Network structure of cerebral cortex shapes functional connectivity on multiple time scales. *Proc. Natl Acad. Soc. USA* **104**, 10240–10245 (2007).

27. van Essen, D. C. et al. The WU–Minn Human Connectome Project: an overview. *NeuroImage* **80**, 62–79 (2013).

28. Glasser, M. F. et al. A multi-modal parcellation of human cerebral cortex. *Nature* **536**, 171–178 (2016).

29. Naze, S., Proix, T., Atasoy, S. & Kozloski, J. R. Robustness of connectome harmonics to local gray matter and long-range white matter connectivity changes: sensitivity analysis of connectome harmonics. *NeuroImage* **224**, 117364 (2021).

30. Robinson, P. A., Loxley, P. N., O'Connor, S. C. & Rennie, C. J. Modal analysis of corticothalamic dynamics, electroencephalographic spectra, and evoked potentials. *Phys. Rev. E* **63**, 041909 (2001).

31. Wingeier, B. M., Nunez, P. L. & Silberstein, R. B. Spherical harmonic decomposition applied to spatial-temporal analysis of human high-density electroencephalogram. *Phys. Rev. E* **64**, 051916 (2001).

32. Barch, D. M. et al. Function in the human connectome: task-fMRI and individual differences in behavior. *NeuroImage* **80**, 169–189 (2013).

33. Gorgolewski, K. J. et al. NeuroVault.Org: a web-based repository for collecting and sharing unthresholded statistical maps of the human brain. *Front. Neuroinform.* https://doi.org/10.3389/fninf.2015.00008 (2015).

34. Lévy, B. Laplace–Beltrami eigenfunctions towards an algorithm that 'understands' geometry. In *Proc. IEEE International Conference on Shape Modeling and Applications 2006 (SMI 2006)* 13 (IEEE, 2006).

35. Robinson, P. A. et al. Prediction of electroencephalographic spectra from neurophysiology. *Phys. Rev. E* **63**, 021903 (2001).

36. Raut, R. V. et al. Global waves synchronize the brain's functional systems with fluctuating arousal. *Sci. Adv.* **7**, eabf2709 (2021).

37. Bolt, T. et al. A parsimonious description of global functional brain organization in three spatiotemporal patterns. *Nat. Neurosci.* **25**, 1093–1103 (2022).

38. Deco, G. et al. How local excitation-inhibition ratio impacts the whole brain dynamics. *J. Neurosci.* **34**, 7886–7898 (2014).

39. Mitra, A., Snyder, A. Z., Blazey, T. & Raichle, M. E. Lag threads organize the brain's intrinsic activity. *Proc. Natl Acad. Sci. USA* **112**, E2235–E2244 (2015).

40. Felleman, D. J. & Van Essen, D. C. Distributed hierarchical processing in the primate cerebral cortex. *Cereb. Cortex* **1**, 1–47 (1991).

41. Chaudhuri, R., Knoblauch, K., Gariel, M. A., Kennedy, H. & Wang, X. J. A large-scale circuit mechanism for hierarchical dynamical processing in the primate cortex. *Neuron* **88**, 419–431 (2015).

42. Goodale, M. A. & Milner, A. D. Separate visual pathways for perception and action. *Trends Neurosci.* **15**, 20–25 (1992).

43. Markov, N. T. et al. Anatomy of hierarchy: feedforward and feedback pathways in macaque visual cortex. *J. Comp. Neurol.* **522**, 225–259 (2014).

44. Hasson, U., Yang, E., Vallines, I., Heeger, D. J. & Rubin, N. A hierarchy of temporal receptive windows in human cortex. *J. Neurosci.* **28**, 2539–2550 (2008).

45. Murray, J. D. et al. A hierarchy of intrinsic timescales across primate cortex. *Nat. Neurosci.* **17**, 1661–1663 (2014).

46. Glasser, M. F. & Van Essen, D. C. Mapping human cortical areas in vivo based on myelin content as revealed by T1- and T2-weighted MRI. *J. Neurosci.* **31**, 11597–11616 (2011).

47. Gao, R., Van den Brink, R. L., Pfeffer, T. & Voytek, B. Neuronal timescales are functionally dynamic and shaped by cortical microarchitecture. *eLife* **9**, e61277 (2020).

48. Tian, Y., Margulies, D. S., Breakspear, M. & Zalesky, A. Topographic organization of the human subcortex unveiled with functional connectivity gradients. *Nat. Neurosci.* **23**, 1421–1432 (2020).

49. Haak, K. V., Marquand, A. F. & Beckmann, C. F. Connectopic mapping with resting-state fMRI. *NeuroImage* **170**, 83–94 (2018).

50. Deco, G., Jirsa, V. K. & McIntosh, A. R. Emerging concepts for the dynamical organization of resting-state activity in the brain. *Nat. Rev. Neurosci.* **12**, 43–56 (2011).

51. Roberts, J. A. et al. Metastable brain waves. *Nat. Commun.* **10**, 1056 (2019).

52. Pang, J. C., Gollo, L. L. & Roberts, J. A. Stochastic synchronization of dynamics on the human connectome. *NeuroImage* **229**, 117738 (2021).

53. Rosen, B. Q. & Halgren, E. An estimation of the absolute number of axons indicates that human cortical areas are sparsely connected. *PLoS Biol.* **20**, e3001575 (2022).

54. van den Heuvel, M. P., Kahn, R. S., Goñi, J. & Sporns, O. High-cost, high-capacity backbone for global brain communication. *Proc. Natl Acad. Sci. USA* **109**, 11372–11377 (2012).

55. Arnatkeviciute, A. et al. Genetic influences on hub connectivity of the human connectome. *Nat. Commun.* **12**, 4237 (2021).

56. Pang, J. C., Rilling, J. K., Roberts, J. A., van den Heuvel, M. P. & Cocchi, L. Evolutionary shaping of human brain dynamics. *eLife* **11**,

e80627 (2022).

57. Aquino, K. M., Fulcher, B. D., Parkes, L., Sabaroedin, K. & Fornito, A. Identifying and removing widespread signal deflections from fMRI data: rethinking the global signal regression problem. *NeuroImage* **212**, 116614 (2020).

58. Gajwani, M. et al. Can hubs of the human connectome be identified consistently with diffusion MRI? Preprint at *bioRxiv* https://doi.org/10.1101/2022.12.21.521366 (2022).

59. Hamid, A. A., Frank, M. J. & Moore, C. I. Wave-like dopamine dynamics as a mechanism for spatiotemporal credit assignment. *Cell* **184**, 2733–2749 (2021).

60. Yousefi, B. & Keilholz, S. Propagating patterns of intrinsic activity along macroscale gradients coordinate functional connections across the whole brain. *NeuroImage* **231**, 117827 (2021).

61. Coalson, T. S., Van Essen, D. C. & Glasser, M. F. The impact of traditional neuroimaging methods on the spatial localization of cortical areas. *Proc. Natl Acad. Sci. USA* **115**, E6356–E6365 (2018).

62. Robinson, P. A. et al. Determination of dynamic brain connectivity via spectral analysis. *Front. Hum. Neurosci.* **15**, 655576 (2021).

63. Majeed, W. et al. Spatiotemporal dynamics of low frequency BOLD fluctuations in rats and humans. *NeuroImage* **54**, 1140–1150 (2011).

64. Matsui, T., Murakami, T. & Ohki, K. Transient neuronal coactivations embedded in globally propagating waves underlie resting-state functional connectivity. *Proc. Natl Acad. Sci. USA* **113**, 6556–6561 (2016).

65. Deco, G. et al. Dynamical consequences of regional heterogeneity in the brain's transcriptional landscape. *Sci. Adv.* **7**, eabf4752 (2021).

66. Henderson, J. A., Aquino, K. M. & Robinson, P. A. Empirical estimation of the eigenmodes of macroscale cortical dynamics: reconciling neural field eigenmodes and resting-state networks. *Neuroimage Rep.* **2**, 100103 (2022).

67. Fischl, B. FreeSurfer. *NeuroImage* **62**, 774–781 (2012).

68. Oldham, S. et al. The efficacy of different preprocessing steps in reducing motion-related confounds in diffusion MRI connectomics. *NeuroImage* **222**, 117252 (2020).

69. Arslan, S. et al. Human brain mapping: a systematic comparison of parcellation methods for the human cerebral cortex. *NeuroImage* **170**, 5–30 (2018).

70. Tokariev, A. et al. Large-scale brain modes reorganize between infant sleep states and carry prognostic information for preterms. *Nat. Commun.* **10**, 2619 (2019).

71. Chen, Y.-C. et al. The individuality of shape asymmetries of the human cerebral cortex. *eLife* **11**, e75056 (2022).

72. Wachinger, C., Golland, P., Kremen, W., Fischl, B. & Reuter, M. BrainPrint: a discriminative characterization of brain morphology. *NeuroImage* **109**, 232–248 (2015).

73. Chavel, I. *Eigenvalues in Riemannian Geometry* 2nd edn (Elsevier, 1984).

74. Seo, S. & Chung, M. K. Laplace–Beltrami eigenfunction expansion of cortical manifolds. In *Proc. International Symposium on Biomedical Imaging* https://doi.org/10.1109/ISBI.2011.5872426 (IEEE, 2011).

75. Reuter, M., Wolter, F. E. & Peinecke, N. Laplace-Beltrami spectra as 'Shape-DNA' of surfaces and solids. *Comput. Aided Des.* **38**, 342–366 (2006).

76. Goscinski, W. J. et al. The multi-modal Australian ScienceS Imaging and Visualization Environment (MASSIVE) high performance computing infrastructure: applications in neuroscience and neuroinformatics research. *Front. Neuroinform.* **8**, 30 (2014).

77. Fischl, B., Sereno, M. I., Tootell, R. B. H. & Dale, A. M. High-resolution intersubject averaging and a coordinate system for the cortical surface. *Hum. Brain Mapp.* **8**, 272–284 (1999).

78. Shuman, D. I., Narang, S. K., Frossard, P., Ortega, A. & Vandergheynst, P. The emerging field of signal processing on graphs. *IEEE Signal Process Mag.* **30**, 83–98 (2013).

79. Courant, R. & Hilbert, D. *Methods of Mathematical Physics* Vol. 1 (Wiley, 1953).

80. Eickhoff, S. B., Yeo, B. T. T. & Genon, S. Imaging-based parcellations of the human brain. *Nat. Rev. Neurosci.* **19**, 672–686 (2018).

81. Mansour L, S., Tian, Y., Yeo, B. T. T., Cropley, V. & Zalesky, A. High-resolution connectomic fingerprints: mapping neural identity and behavior. *NeuroImage* **229**, 117695 (2021).

82. Schaefer, A. et al. Local-global parcellation of the human cerebral cortex from intrinsic functional connectivity mri. *Cereb. Cortex* **28**, 3095–3114 (2018).

83. Rué-Queralt, J. et al. The connectome spectrum as a canonical basis for a sparse representation of fast brain activity. *NeuroImage* **244**, 118611 (2021).

84. Tournier, J.-D. et al. MRtrix3: a fast, flexible and open software framework for medical image processing and visualisation. *NeuroImage* **202**, 116137 (2019).

85. Theodoni, P. et al. Structural attributes and principles of the neocortical connectome in the marmoset monkey. *Cereb. Cortex* **32**, 15–28 (2022).

86. Abraham, A. et al. Machine learning for neuroimaging with scikit-learn. *Front. Neuroinformat.* **8**, 14 (2014).

87. Robinson, P. A. Interrelating anatomical, effective, and functional brain connectivity using propagators and neural field theory. *Phys. Rev. E* **85**, 011912 (2012).

88. Sanz-Leon, P. et al. The virtual brain: a simulator of primate brain network dynamics. *Front. Neuroinformat.* **7**, 10 (2013).

89. Breakspear, M., Heitmann, S. & Daffertshofer, A. Generative models of cortical oscillations: neurobiological implications of the Kuramoto model. *Front. Hum. Neurosci.* **4**, 190 (2010).

90. Wilson, H. R. & Cowan, J. D. Excitatory and inhibitory interactions in localized populations of model neurons. *Biophys. J.* **12**, 1–24 (1972).

91. Demirtaş, M. et al. Hierarchical heterogeneity across human cortex shapes large-scale neural dynamics. *Neuron* **101**, 1181–1194 (2019).

92. Aquino, K. M. et al. On the intersection between data quality and dynamical modelling of large-scale fMRI signals. *NeuroImage* **256**, 119051 (2022).

93. Behrens, T. E. J. et al. Characterization and propagation of uncertainty in diffusion-weighted MR imaging. *Magn. Reson. Med.* **50**, 1077–1088 (2003).

94. Deco, G., Jirsa, V. K., Robinson, P. A., Breakspear, M. & Friston, K. The dynamic brain: from spiking neurons to neural masses and cortical fields. *PLoS Comput. Biol.* **4**, e1000092 (2008).

95. Spiegler, A. & Jirsa, V. Systematic approximations of neural fields through networks of neural masses in the virtual brain. *NeuroImage* **83**, 704–725 (2013).

96. Stephan, K. E., Weiskopf, N., Drysdale, P. M., Robinson, P. A. & Friston, K. J. Comparing hemodynamic models with DCM.

NeuroImage **38**, 387–401 (2007).

97. Aquino, K. M., Schira, M. M., Robinson, P. A., Drysdale, P. M. & Breakspear, M. Hemodynamic traveling waves in human visual cortex. *PLoS Comput. Biol.* **8**, e1002435 (2012).

98. Pang, J. C., Robinson, P. A., Aquino, K. M. & Vasan, N. Effects of astrocytic dynamics on spatiotemporal hemodynamics: modeling and enhanced data analysis. *NeuroImage* **147**, 994–1005 (2017).

99. Pang, J. C., Aquino, K. M., Robinson, P. A., Lacy, T. C. & Schira, M. M. Biophysically based method to deconvolve spatiotemporal neurovascular signals from fMRI data. *J. Neurosci. Methods* **308**, 6–20 (2018).

100. Deco, G., Jirsa, V., McIntosh, A. R., Sporns, O. & Kötter, R. Key role of coupling, delay, and noise in resting brain fluctuations. *Proc. Natl Acad. Sci. USA* **106**, 10302–10307 (2009).

101. Cabral, J., Kringelbach, M. L. & Deco, G. Exploring the network dynamics underlying brain activity during rest. *Prog. Neurobiol.* **114**, 102–131 (2014).

102. Heinzle, J., Koopmans, P. J., den Ouden, H. E. M., Raman, S. & Stephan, K. E. A hemodynamic model for layered BOLD signals. *NeuroImage* **125**, 556–570 (2016).

103. Sanz-Leon, P. et al. NFTsim: theory and simulation of multiscale neural field dynamics. *PLoS Comput. Biol.* **14**, e1006387 (2018).

104. Deco, G., Kringelbach, M. L., Jirsa, V. K. & Ritter, P. The dynamics of resting fluctuations in the brain: metastability and its dynamical cortical core. *Sci. Rep.* **7**, 3095 (2017).

105. Glerean, E., Salmi, J., Lahnakoski, J. M., Jääskeläinen, I. P. & Sams, M. Functional magnetic resonance imaging phase synchronization as a measure of dynamic functional connectivity. *Brain Connect.* **2**, 91–101 (2012).

106. Pang, J. C. & Robinson, P. A. Power spectrum of resting-state blood-oxygen-level-dependent signal. *Phys. Rev. E* **100**, 022418 (2019).

107. Mitra, A., Snyder, A. Z., Hacker, C. D. & Raichle, M. E. Lag structure in resting-state fMRI. *J. Neurophysiol.* **111**, 2374–2391 (2014).

108. Cabral, J., Kringelbach, M. L. & Deco, G. Functional connectivity dynamically evolves on multiple time-scales over a static structural connectome: models and mechanisms. *NeuroImage* **160**, 84–96 (2017).

109. Kashyap, A. & Keilholz, S. Dynamic properties of simulated brain network models and empirical resting-state data. *Netw. Neurosci.* **3**, 405–426 (2019).

110. Burt, J. B. et al. Hierarchy of transcriptomic specialization across human cortex captured by structural neuroimaging topography. *Nat. Neurosci.* **21**, 1251–1259 (2018).

111. Váša, F. et al. Adolescent tuning of association cortex in human structural brain networks. *Cereb. Cortex* **28**, 281–294 (2018).

112. Alexander-Bloch, A. F. et al. On testing for spatial correspondence between maps of human brain structure and function. *NeuroImage* **178**, 540–551 (2018).

113. Uğurbil, K. et al. Pushing spatial and temporal resolution for functional and diffusion MRI in the Human Connectome Project. *NeuroImage* **80**, 80–104 (2013).

114. Vos de Wael, R. et al. Anatomical and microstructural determinants of hippocampal subfield functional connectome embedding. *Proc. Natl Acad. Sci. USA* **115**, 10154–10159 (2018).

115. Yang, S. et al. The thalamic functional gradient and its relationship to structural basis and cognitive relevance. *NeuroImage* **218**, 116960 (2020).

116. Oldehinkel, M. et al. Mapping dopaminergic projections in the human brain with resting-state fMRI. *eLife* **11**, e71846 (2022).

117. Cohen, A. L. et al. Defining functional areas in individual human brains using resting functional connectivity MRI. *NeuroImage* **41**, 45–57 (2008).

118. Belkin, M. & Niyogi, P. Laplacian eigenmaps for dimensionality reduction and data representation. *Neural Comput.* **15**, 1373–1396 (2003).

119. Margulies, D. S. et al. Situating the default-mode network along a principal gradient of macroscale cortical organization. *Proc. Natl Acad. Sci. USA* **113**, 12574–12579 (2016).

Acknowledgements

We thank S. Mansour L for providing access to high-resolution connectome data and K. Haak for assistance with connectopic mapping. HCP data were provided by the HCP, Wu–Minn Consortium (principal Investigators D. Van Essen and K. Ugurbil; no. 1U54MH091657) funded by the 16 NIH Institutes and centres that support the NIH Blueprint for Neuroscience Research, and by the McDonnell Center for Systems Neuroscience at Washington University. This work was supported by the MASSIVE HPC facility (www.massive.org.au), Sylvia and Charles Viertel Foundation grant no. 2017042 to A.F., National Health and Medical Research Council grant nos. 1197431 and 1146292 to A.F., Australian Research Council grant nos. FL220100184 and DP200103509 to A.F., National Health and Medical Research Council grant no. 2008612 to M.B., Australian Research Council Laureate Fellowship no. FL140100025 to P.A.R. and Australian Research Council Center of Excellence no. CE140100007 to P.A.R.

Author information

Author notes

1. These authors contributed equally: James C. Pang, Kevin M. Aquino

Authors and Affiliations

1. The Turner Institute for Brain and Mental Health, School of Psychological Sciences and Monash Biomedical Imaging, Monash University, Clayton, Victoria, Australia

   James C. Pang & Alex Fornito

2. School of Physics, University of Sydney, Camperdown, New South Wales, Australia

   Kevin M. Aquino, Peter A. Robinson & Ben D. Fulcher

3. BrainKey Inc., San Francisco, CA, USA

   Kevin M. Aquino

4. Donders Institute for Brain, Cognition, and Behaviour, Radboud University Medical Centre, Nijmegen, the Netherlands

   Marianne Oldehinkel

5. School of Psychological Sciences, College of Engineering, Science and the Environment, University of Newcastle, Callaghan, New South Wales, Australia

   Michael Breakspear

6. School of Medicine and Public Health, College of Health, Medicine and Wellbeing, University of Newcastle, Callaghan, New South Wales, Australia

Michael Breakspear

Contributions

K.M.A. and A.F. conceptualized the study. J.C.P., K.M.A., A.F. and M.O. designed the methodology. J.C.P., K.M.A. and A.F. performed the investigation and administered the project. J.C.P. and K.M.A. developed visualizations. A.F. acquired funding. A.F., P.A.R., B.D.F. and M.B. supervised the project. J.C.P. and A.F. wrote the original draft. All authors reviewed and edited the final manuscript.

Corresponding author

Correspondence to James C. Pang.

Ethics declarations

Competing interests

K.M.A. is a scientific advisor and shareholder in BrainKey Inc., a medical image analysis software company. The other authors declare no competing interests.

Peer review

Peer review information

Nature thanks R. Todd Constable, Shella Keilholz, Maria Giulia Preti and Enzo Tagliazucchi for their contribution to the peer review of this work. Peer reviewer reports are available.

Additional information

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Extended data figures and tables

Extended Data Fig. 1 Eigenmode basis sets.

The basis sets from left to right are geometric eigenmodes, connectome eigenmodes, connectome eigenmodes using a connectivity matrix matching the density used by the exponential distance rule (EDR) eigenmodes, and EDR eigenmodes. Negative–zero–positive values are coloured as blue–white–red.

Extended Data Fig. 2 Reconstruction accuracy of resting-state FC achieved by different basis sets for different parcellation resolution.

The basis sets are geometric eigenmodes, connectome eigenmodes, connectome eigenmodes using a connectivity matrix with density matched to that used by the EDR eigenmodes, and EDR eigenmodes. Schaefer100, Schaefer200, Glasser360, Schaefer400, Schaefer600, Schaefer800, and Schaefer1000 has 100, 200, 360, 400, 600, 800, and 1000 parcels, respectively, across both hemispheres.

Extended Data Fig. 3 Difference in reconstruction accuracy of all 47 HCP task-contrast maps achieved by geometric eigenmodes and connectome eigenmodes for different parcellation resolution.

Each row represents a different task contrast, which have been grouped here by broad types (Supplementary Information 2.1). wm = working memory. Red indicates superior performance for geometric eigenmodes. Schaefer100, Schaefer200, Glasser360, Schaefer400, Schaefer600, Schaefer800, and Schaefer1000 has 100, 200, 360, 400, 600, 800, and 1000 parcels, respectively, across both hemispheres.

Extended Data Fig. 4 Reconstruction accuracy achieved by geometric eigenmodes and PCA.

(**a**) Reconstruction accuracy of resting-state FC. (**b**) Comparison of the reconstruction accuracy of all 47 HCP task-contrast maps, which have been grouped here by broad types (Supplementary Information 2.1). wm = working memory. Each row represents a different task contrast. Red indicates superior performance for geometric eigenmodes. The asterisk denotes the contrast (that is, the seven key HCP task contrasts) within the relevant task used to train the PCA.

Extended Data Fig. 5 Comparison of geometric eigenmodes and Fourier basis sets.

a, Spatial maps of modes 1, 2, 3, 4, 10, 50 and 100 of six different Fourier basis sets with unit coefficients. The terms reg and irreg mean that the spatial wavelengths of the modes in the x-, y-, and z-directions are spaced in regular and irregular increments, respectively. See Supplementary Information 6 for details. **b**, Reconstruction accuracy of seven key HCP task-contrast maps and resting-state FC. See Supplementary Information 2.1 for details about the contrast maps. wm = working memory.

Extended Data Fig. 6 Classical neuroimaging approach of thresholding statistical maps.

a, Simple one-dimensional example of how different thresholds only capture focal clusters of activations and ignore the underlying structured pattern of activations. **b**, Spatially embedded demonstration of the concept depicted in **a** using unthresholded and binarized thresholded maps.

Extended Data Fig. 7 Normalized power spectrum of each of the seven key HCP task-contrast maps.

See Supplementary Information 2.1 for details about the contrast maps. wm = working memory.

Extended Data Fig. 8 Power spectra of empirical task-activation maps and surrogate maps.

a, Normalized mean power spectra of 47 HCP task-contrast maps (top) and 10,000 contrast maps from the NeuroVault database (bottom). The coloured lines correspond to power spectra of surrogate data following the application of spatial smoothing filters with varying full-width at half-maximum (FWHM). **b**, Average mean square logarithmic error (MSLE) as a function of FWHM between normalized mean power spectra of HCP and NeuroVault contrast maps and smoothed surrogate data. **c**, MSLE separately obtained between the power spectra of each of the 47 HCP and 10,000 NeuroVault contrast maps and the smoothed surrogate data. Each row represents a different task-contrast map. The lines correspond to the FWHM where MSLE is minimum for each map.

Extended Data Fig. 9 Comparison of the wave and neural mass models in capturing time-lagged properties of fMRI data.

a, Time-delay matrices from empirical data, simulated data using the wave model and simulated data using the neural mass model of the left hemisphere. Negative–zero–positive values are coloured as blue–white–red. **b**, Mean lags from the matrices in **a** (mean of each column) projected on the cortical surface. Negative–zero–positive values are coloured as blue–white–red. The scatter plots show the relationship of mean lags from empirical data and simulated data from the two models for 180 brain regions. The red line represents a linear fit with Pearson correlation coefficient *r* and one-sided spin-test *p*-value, p_{spin}, estimated from 10,000 permutations. **c**, Similar to **b** but on the first principal component (PC1) of the matrices in **a**. The number above the surfaces (var) corresponds to the variance explained by the PC. **d**, Similar to **c** but on the second PC (PC2) of the matrices in **a**.

Extended Data Fig. 10 Optimization of the wave model.

The model is trained on 125 HCP individuals to find the optimal value of the parameter r_s (in mm). Optimization performance compares data and model FC based on the following metrics: edge FC correlation, node FC correlation and FCD KS statistic. Higher edge FC correlation, higher node FC correlation and lower FCD KS statistic correspond to better model fit.

We take $r_s = 28.9$ mm as the optimal parameter as it leads to the minimum FCD KS statistic value.

Supplementary information

Supplementary Information

This file contains Supplementary Information, including Figs. 1–11, Tables 1–3 and additional references.

Reporting Summary

Peer Review File

Supplementary Video 1

Stimulus-evoked activity. Wave propagation of activity after 1 ms stimulation of V1 for $t = 1–2$ ms.

Rights and permissions

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by/4.0/.

Reprints and Permissions

About this article

Cite this article

Pang, J.C., Aquino, K.M., Oldehinkel, M. *et al.* Geometric constraints on human brain function. *Nature* **618**, 566–574 (2023). https://doi.org/10.1038/s41586-023-06098-1

- Received: 14 October 2022
- Accepted: 18 April 2023
- Published: 31 May 2023
- Issue Date: 15 June 2023
- DOI: https://doi.org/10.1038/s41586-023-06098-1

Share this article

Anyone you share the following link with will be able to read this content:

Get shareable link

Sorry, a shareable link is not currently available for this article.

Copy to clipboard

Provided by the Springer Nature SharedIt content-sharing initiative

The human brain’s characteristic wrinkles help to drive how it works

- Davide Castelvecchi

Nature News 31 May 2023

The geometry of the human brain shapes its function

Nature Research Briefing 31 May 2023

This article was downloaded by **calibre** from https://www.nature.com/articles/s41586-023-06098-1

| Section menu | Main menu |

- Article
- Published: 31 May 2023

The effects of cash transfers on adult and child mortality in low- and middle-income countries

- Aaron Richterman ,
- Christophe Millien,
- Elizabeth F. Bair,
- Gregory Jerome,
- Jean Christophe Dimitri Suffrin,
- Jere R. Behrman &
- …
- Harsha Thirumurthy

Show authors

Nature **volume 618**, pages 575–582 (2023)

- 25k Accesses
- 1 Citations
- 463 Altmetric
- Metrics details

Abstract

Poverty is an important social determinant of health that is associated with increased risk of death[1,2,3,4,5]. Cash transfer programmes provide non-

contributory monetary transfers to individuals or households, with or without behavioural conditions such as children's school attendance[6,7]. Over recent decades, cash transfer programmes have emerged as central components of poverty reduction strategies of many governments in low- and middle-income countries[6,7]. The effects of these programmes on adult and child mortality rates remains an important gap in the literature, however, with existing evidence limited to a few specific conditional cash transfer programmes, primarily in Latin America[8,9,10,11,12,13,14]. Here we evaluated the effects of large-scale, government-led cash transfer programmes on all-cause adult and child mortality using individual-level longitudinal mortality datasets from many low- and middle-income countries. We found that cash transfer programmes were associated with significant reductions in mortality among children under five years of age and women. Secondary heterogeneity analyses suggested similar effects for conditional and unconditional programmes, and larger effects for programmes that covered a larger share of the population and provided larger transfer amounts, and in countries with lower health expenditures, lower baseline life expectancy, and higher perceived regulatory quality. Our findings support the use of anti-poverty programmes such as cash transfers, which many countries have introduced or expanded during the COVID-19 pandemic, to improve population health.

Main

Poverty has long been recognized as an important social determinant of health. Poverty can negatively influence health outcomes through numerous, often interconnected pathways—food insecurity, access to and quality of healthcare, housing stability, neighbourhood safety, occupational risk, educational attainment, health behaviours, and social well-being, among others[15,16,17,18,19]. Consequently, living in poverty has been closely linked to a decrease in life expectancies, with a greater risk of mortality among both adults and children[1,2,3,4,5].

Despite many years of progress, nearly 10% of the world's population lived on less than US$1.90 per day (extreme poverty) in 2018, and more than 40% lived on less than US$5.50 per day[20] (upper middle-income poverty line). The COVID-19 pandemic has markedly worsened these figures—an

estimated 97 million more people lived in extreme poverty in 2020 (a 12% increase) and additional increases were seen in low-income countries in 2021[21]. These enduring pandemic-related effects make the assessment and implementation of evidence-based strategies to combat poverty and improve health an even more urgent priority.

Over the past two decades, more than 100 low- and middle-income countries (LMICs) have introduced cash transfer programmes as components of their poverty reduction and social protection strategies[6]. Cash transfer programmes are defined as those that provide non-contributory monetary transfers to individuals or households. They include unconditional transfers (more common in sub-Saharan Africa), conditional transfers (more common in Latin America), public pensions and enterprise grants (money provided to support income-generating activities).

Cash transfer programmes have become even more common during the COVID-19 pandemic. A World Bank report in February 2022 identified 962 cash transfer programmes in 203 countries—672 of these were newly introduced during the pandemic[7]. Indeed, it is estimated that cash transfers were distributed to 1.36 billion people—17% of the world's population—during the pandemic period[22].

Large-scale, government-run cash transfer programmes have been successful in reducing poverty and improving economic autonomy, school attendance, child nutrition, women's empowerment and health-service use among beneficiaries[23,24]. A few studies have also documented population-wide effects such as greater economic activity in communities where beneficiaries reside[25], and—in the case of infectious diseases such as HIV—reduced new infections following the introduction of cash transfer programmes[26]. The improvements seen with cash transfers could be driven by the removal of economic and psychological barriers of poverty as a result of receiving cash transfers, as well as spillover effects on non-beneficiaries[27,28,29,30,31,32].

Despite the large body of literature on the effects of cash transfer programmes on various outcomes, there is limited evidence about the effect of such programmes on overall, population-level mortality rates, particularly outside of a few conditional cash transfer programmes in Latin America.

Several municipal-level analyses have shown a decline in infant mortality associated with the *Bolsa Familia* programme in Brazil[8,9,10]. An individual-level analysis found 17% decreased odds of mortality among children aged less than 5 years who were beneficiaries of *Bolsa Familia*, with stronger associations for children from the poorest communities[11]. Other single-country municipal-level analyses have suggested reductions in infant mortality associated with conditional cash transfer programmes in Mexico, Ecuador and India[33,34,35].

There are even fewer studies of relationships between cash transfer programmes and adult mortality rates. Evaluations of the Mexican conditional cash transfer programme *Oportunidades* found an 11% decline in maternal mortality and a 4% decline in overall mortality in regions where the programme had been phased in[12,13]. A municipal-level study of *Bolsa Familia* similarly found a 10–20% reduction in maternal mortality[14]. In an analysis of 42 countries, we found that cash transfer programmes were associated with population-wide reductions in AIDS-related deaths that grew larger over time[26]. Notably, however, most randomized and non-randomized evaluations of cash transfers have lacked adequate sample sizes or study durations to detect differences in adult or child mortality. The design of most country-specific evaluations is typically also focused on estimating programme effects on beneficiaries rather than the entire population. Unlike large-scale, multinational evaluations of major health aid programmes such as the US President's Emergency Plan for AIDS Relief[36,37] (PEPFAR), no such multinational studies assess the effectiveness of cash transfer programmes in reducing population-level adult and child mortality rates.

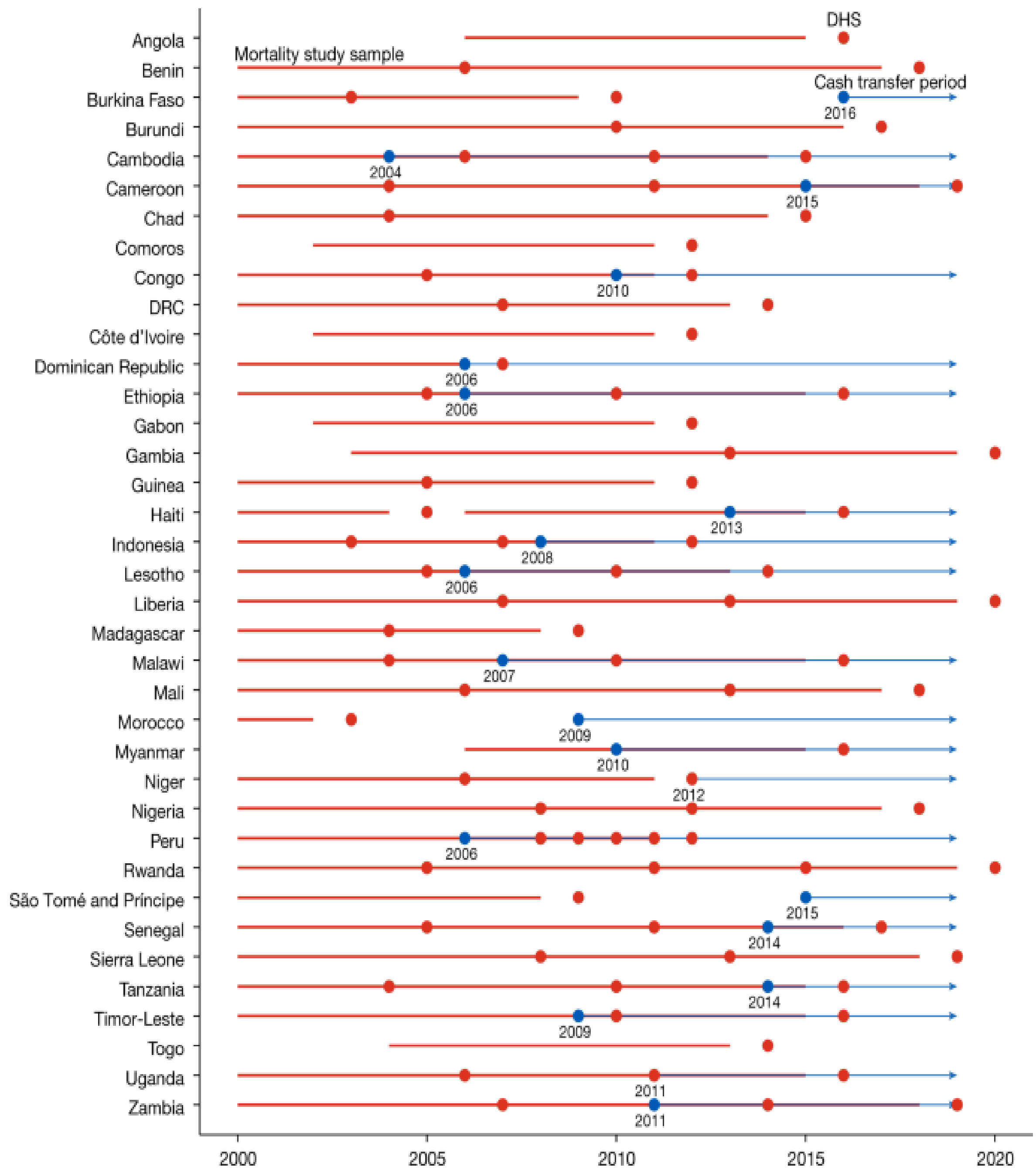
Given the growing popularity of cash transfer programmes, evaluating their overall effects on adult and child mortality rates remains an important and policy-relevant gap in the literature. To close this gap, we used multinational longitudinal data generated from sibling and birth histories collected in national household surveys to evaluate the effects of cash transfer programmes on adult and child mortality among more than seven million people from 2000 to 2019. We used a difference-in-differences approach, a quasi-experimental technique that can be used to estimate causal effects from observational data by comparing the differences in outcomes between intervention and comparison groups during pre-intervention and post-

intervention periods, under an assumption of parallel trends (that is, that in the absence of cash transfer programmes, trends in outcomes would be similar in intervention and comparison countries). Our primary finding was that these programmes were associated with significant mortality reductions among women and children aged less than 5 years, indicating the important role that these anti-poverty initiatives have had in promoting population health over the last 20 years.

Cash transfer programmes and mortality datasets

There were 37 LMICs included in our analysis (see Methods, 'Mortality data' and 'Cash transfer programme data' for selection criteria and Supplementary Table 1 for excluded countries)—29 in sub-Saharan Africa, 3 in Latin America and the Caribbean, 4 in the Asia–Pacific region, and 1 in northern Africa (Fig. 1). Sixteen countries introduced large-scale cash transfer programme(s) during the study period and had mortality data available during their respective cash transfer periods (see Methods, 'Cash transfer programme data' and 'Statistical analysis' for how we identified programmes and defined the cash transfer programme exposure, and Extended Data Fig. 1 for the country inclusion flow diagram).

Fig. 1: Study timeline.



The study period (2000–2019) is along the *x*-axis, included countries (*n* = 37) are listed on the *y*-axis, red points represent national demographic and health surveys (*n* = 84), red lines represent corresponding years with included mortality data generated from sibling and birth histories, blue points represent the first complete year of cash transfer programme(s) covering more than 5% of the impoverished population (*n* = 20 total; *n* = 16 with mortality data during the cash transfer period), and blue lines represent

the cash transfer period. DHS, demographic and health survey; DRC, Democratic Republic of the Congo.

Within these 16 'intervention countries', there were 29 total cash transfer programmes, 14 (48%) of which were unconditional (Supplementary Tables 2 and 3 show programme-specific details). Intervention countries had a median most recent impoverished population coverage of 27% (interquartile range 16–100%), with a median most recent maximum transfer amount per beneficiary equating to 10% of per capita gross domestic product (GDP) (interquartile range 6.25–13.25%). Six countries had high (above-median) coverage with high (above-median) maximum transfer amounts, two had high coverage with low maximum transfer amounts, two had low coverage with high maximum transfer amounts, and six had low coverage with low maximum transfer amounts.

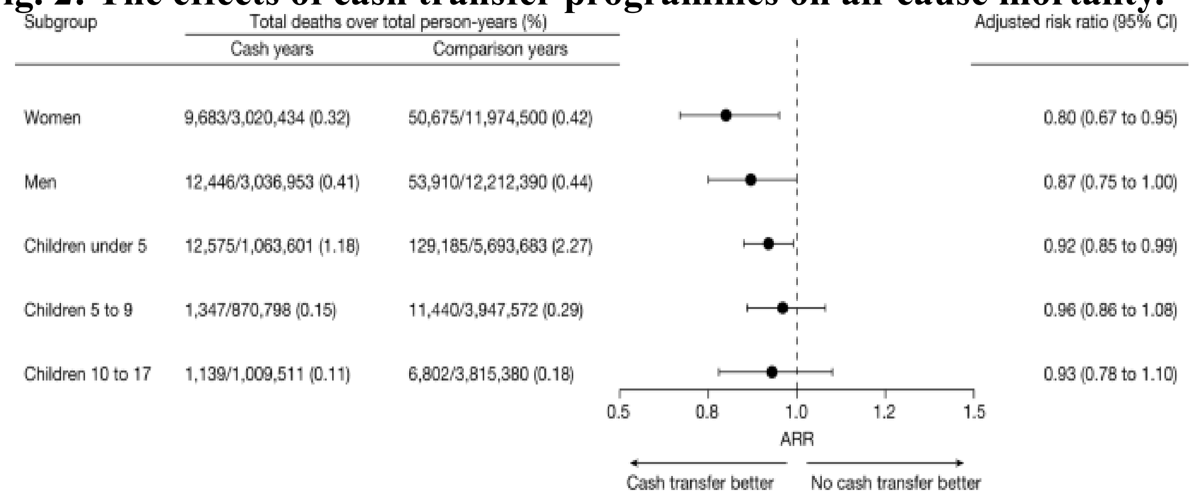
There were 4,325,484 people included in the adult dataset, with a total of 30,244,277 person-years (6,057,387 (20%) during intervention years) and 126,714 deaths (42 per 10,000 person-years) (Supplementary Table 4; see Methods, 'Mortality data' for details about the generation of the mortality datasets). There were 2,867,940 people included in the child dataset, with a total of 16,400,545 person-years (2,943,910 (18%) during intervention years) and 162,488 deaths (99 per 10,000 person-years) (Supplementary Table 5). For both datasets, comparison person-years had lower GDP per capita, lower percentiles for the World Bank Worldwide Governance Indicators, and a greater proportion of person-years from sub-Saharan Africa (Extended Data Tables 1 and 2).

Effects of cash transfers on mortality

In our primary difference-in-differences analyses, cash transfer programmes were associated with reductions in mortality among women (adult female individuals at least 18 years of age) (adjusted risk ratio (ARR) 0.80, 95% confidence interval 0.67–0.95) and children aged less than 5 years (ARR 0.92, 95% confidence interval 0.85–0.99) (Fig. 2 and Supplementary Tables 6–10; see Methods, 'Statistical analysis' for additional details about the models). These reductions are at the higher end of the range of estimates from single-country studies of specific cash transfer

programmes[8,9,10,11,12,13,14,33,34,35]. There were no associations between cash transfer programmes and mortality among men (adult male individuals at least 18 years of age) (ARR 0.87, 95% confidence interval 0.75–1.00), children aged 5–9 years (ARR 0.96, 95% confidence interval 0.86–1.08) or children aged 10–17 years (ARR 0.93, 95% confidence interval 0.78–1.10).

Fig. 2: The effects of cash transfer programmes on all-cause mortality.

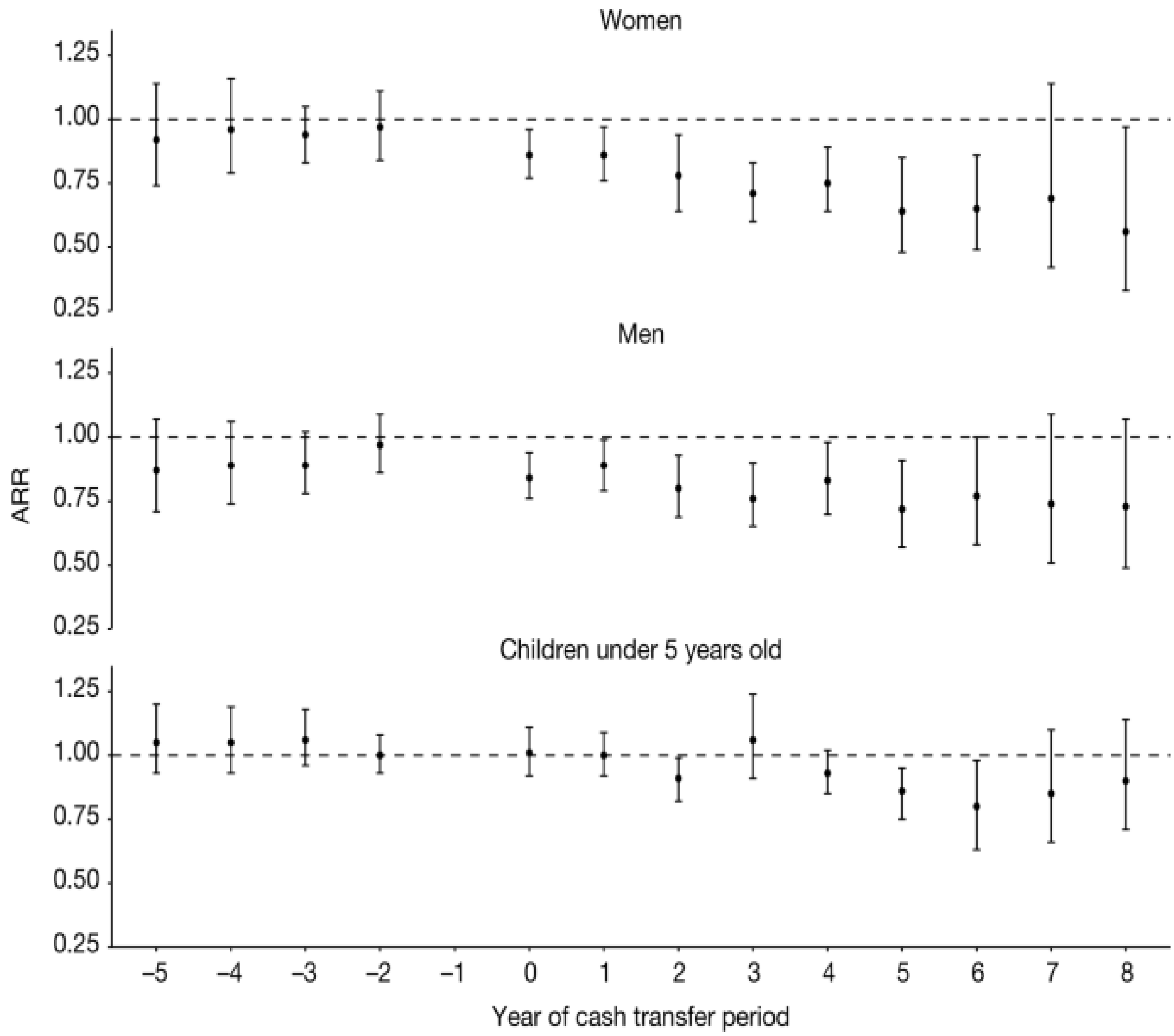


Forest plot showing the fully adjusted overall associations between cash transfer programmes and mortality among women ($n = 14,994,934$ person-years), men ($n = 15,249,343$ person-years) and children aged less than 5 years ($n = 6,757,284$ person-years), 5 to 9 years ($n = 4,818,370$ person-years) and 10 to 17 years ($n = 4,824,891$ person-years). Effect estimates are ARRs and error bars represent 95% confidence intervals. Estimates were generated using multivariable modified Poisson models with country and year fixed effects, country-level covariates (GDP per capita, PEPFAR funding budgeted, and three Worldwide Governance Indicators: control of corruption, political stability and absence of violence, and voice and accountability), and individual-level covariates (age and rural or urban setting in all models; sex, age of mother and birth order in child analyses). We used robust standard errors clustered at the country level. CI, confidence interval.

We next assessed the temporal patterns in relationships between cash transfer programmes and mortality by creating a series of binary indicators for each year before and after each cash transfer period began. Consistent with our primary analyses, fully adjusted models showed that significant

reductions in mortality among adult women and children aged less than 5 years occurred within 2 years of programme introduction (Fig. 3), with even larger reductions detected over time among women. Temporal analyses also suggested reductions in mortality among men over time (Fig. 3). There was no evidence of associations between cash transfer programmes and mortality over time among children aged 5–9 years or 10–17 years (Extended Data Fig. 2).

Fig. 3: The effects of cash transfer programmes on all-cause mortality over time.



Temporal plots showing the associations between cash transfer programmes and mortality as a function of the year of the cash transfer period. Effect estimates are ARRs and error bars show 95% confidence intervals. Estimates were generated using multivariable modified Poisson models with

country and year fixed effects, country-level covariates (GDP per capita, PEPFAR funding budgeted and three Worldwide Governance Indicators: control of corruption, political stability and absence of violence, and voice and accountability), and individual-level covariates (age and rural or urban setting in all models; sex, age of mother and birth order in child analyses). We used robust standard errors clustered at the country level. Top, estimates for women ($n = 14,994,934$ person-years). Middle, estimates for men ($n = 15,249,343$ person-years). Bottom, estimates for children aged less than 5 years ($n = 6,757,284$ person-years).

We also used these temporal plots to show that there were no differential pre-trends in the years before the introduction of cash transfer programmes (Fig. 3 and Extended Data Fig. 2). The parallel trends assumption was further supported by regression models showing that trends in mortality rates were similar between intervention and comparison countries before the introduction of cash transfers (see Methods, 'Statistical analysis' for details of these models, and Supplementary Table 11).

Secondary heterogeneity analyses

We then explored the heterogeneity of the effects of cash transfer programmes on mortality through subgroup analyses based on individual characteristics, cash transfer programme design features and country characteristics (Tables 1 and 2 show subgroups for women; men and children are shown in Extended Data Figs. 3–6). Although these subgroup analyses should be considered exploratory in the setting of multiple comparisons, there were several notable findings.

Table 1 Heterogeneity analyses for women—individual factors
Table 2 Heterogeneity analyses for women—programme and country factors

There was a significant reduction in mortality among men aged 18–40 years (ARR 0.86, 95% confidence interval 0.77–0.96) (Extended Data Fig. 2), with a possible mortality reduction among some men also supported by findings from our temporal analyses. Among women, there were reductions in both pregnancy-related deaths (ARR 0.74, 95% confidence interval 0.61–

0.91) and non-pregnancy-related deaths (ARR 0.81, 95% confidence interval 0.68–0.94).

There were no apparent differences between the effects of unconditional and conditional transfers on mortality. At a minimum this provides reassurance that the mortality benefits of cash transfers were not limited to conditional transfers, which have been the focus of the few country-specific studies that evaluated the effects of cash transfers on mortality[8,9,10,11,12,13,14,33,34,35]. Conditional cash transfer programmes typically incentivize behaviours surrounding nutrition, education or health services use (commonly focused on children), whereas unconditional cash transfer programmes tend to be more direct anti-poverty approaches, have fewer administrative costs, and are more widely used in sub-Saharan Africa.

We also found that programmes with higher coverage and larger cash transfer amounts were associated with the largest reductions in mortality, with these types of programmes being associated with significant reductions among women (ARR 0.70, 95% confidence interval 0.62–0.79), men (ARR 0.77, 95% confidence interval 0.71–0.84), children aged less than 5 years (ARR 0.86, 95% confidence interval 0.81–0.93) and children aged 10–17 years (ARR 0.80, 95% confidence interval 0.65–0.97), but not for children aged 5–9 years (ARR 0.94, 95% confidence interval 0.83–1.07). This finding further supports a causal relationship between cash transfer programmes and risks of death. It also indicates that programmes with lower coverage or transfer amounts may be less effective or ineffective in reducing population-level mortality rates, although the confidence intervals in these lower-coverage, lower-amount groups were generally too wide to draw firm conclusions.

Countries with higher regulatory quality ratings within the Worldwide Governance Indicators generally showed greater reductions in mortality, with significant reductions seen among women (ARR 0.71, 95% confidence interval 0.63–0.80), men (ARR 0.80, 95% confidence interval 0.73–0.88) and children aged less than 5 years (ARR 0.89, 95% confidence interval 0.83–0.94). Findings related to voice and accountability ratings were less intuitive, with significant mortality reductions seen in countries with lower ratings among women (ARR 0.74, 95% confidence interval 0.66–0.85), men (ARR 0.81, 95% confidence interval 0.73–0.90) and children aged less than

5 years (ARR 0.89, 95% confidence interval 0.83–0.95), but not in countries with higher ratings. Subgroup analyses based on these indicators are therefore of unclear overall significance and should be interpreted with caution.

We also found greater effect sizes in countries with lower health expenditures per capita and in those with lower life expectancies (a finding that has been noted elsewhere[10]), indicating that people living in countries with little healthcare infrastructure or substantial public health challenges may especially benefit from cash transfer programmes.

Stratification by region showed a stronger association between cash transfers and mortality among adult women in sub-Saharan Africa (ARR 0.77, 95% confidence interval 0.62–0.95) relative to outside sub-Saharan Africa (ARR 0.93, 95% confidence interval 0.79–1.11).

Despite the heterogeneities noted above, country-specific estimates for the 16 individual intervention countries were largely similar to our primary analyses, indicating benefits of cash transfer programmes across broadly diverse contexts (Extended Data Fig. 7). There were a few noteworthy exceptions—the Dominican Republic (ARR 1.20, 95% confidence interval 0.85–1.68), Indonesia (ARR 1.04, 95% confidence interval 0.82–1.34) and Lesotho (ARR 1.25, 95% confidence interval 1.11–1.41). The Dominican Republic had only one year of mortality data available post-intervention, and Lesotho was the only intervention country in our dataset that had a cash transfer period during which only older adults were targeted by cash transfer programmes. Country-specific estimates may also be more vulnerable to an important confounding factor specific to that country, such as a positive or negative shock or policy changes other than the cash transfer programme occurring approximately simultaneously with its introduction. We therefore caution against placing too much weight on estimates for any single country.

Other sensitivity analyses

Our findings were generally robust to a variety of additional sensitivity analyses (detailed in Methods, ‘Statistical analysis’). Fully adjusted logistic regression models (rather than modified Poisson models) yielded identical

results for all outcomes to two decimal places. Fully adjusted linear models were consistent with those from our primary analyses except that there were overall associations between cash transfer programmes and mortality among men, and there were no longer any associations among children aged less than 5 years (Supplementary Table 12). Recent advances in difference-in-differences analyses with variation in intervention timing have shown that estimates may be biased, particularly if there is heterogeneity in intervention effects over time[38,39,40]. Use of an alternate fully adjusted linear estimator that is not vulnerable to this bias showed highly similar results to standard fully adjusted linear models, which provides reassurance that bias resulting from heterogeneous intervention effects over time is minimal[41] (Supplementary Table 13). This bias also tends to be influenced by later country-years during the intervention period, and excluding intervention years after year five did not substantially influence our effect estimates—although, as with some other modelling approaches, there were now overall associations between cash transfer programmes and mortality among men, and there were no longer any associations among children aged less than five years (Supplementary Table 14). Repeating the adult female analysis with the exclusion of individual countries did not reveal possible outlier countries (Supplementary Table 15). The addition of the survey respondent's wealth quintile and educational attainment to our primary models resulted in minimal changes to our estimates (Supplementary Table 16).

Discussion and conclusions

Although our analytic approach has previously been used to evaluate the relationship between health aid programmes and mortality[36,37], to our knowledge, this is the first study to use it to examine the effects of government-led anti-poverty programmes on population-level mortality rates. Our results were consistent with previous single-country studies predominantly of conditional cash transfer programmes in Latin America[8,9,10,11,12,13,14,33,34,35], although notably we focus on many LMICs outside of Latin America that have higher underlying poverty and mortality. We also study effects on the entire populations rather than on beneficiaries alone. The results are also consistent with our previous multi-country study showing associations between cash transfer programmes and reductions in

AIDS-related deaths in a subset of countries with generalized HIV epidemics[26].

The largest and most convincing mortality reductions were among women. This adds to previous evidence that cash transfers may disproportionately benefit women, or be more effective when women are the primary beneficiaries[23,42,43,44,45]. Reflecting this, many of the cash programmes that we identified either targeted women directly or were designed in ways that favoured women (for example, minimum age-based eligibility will tend to benefit women, who live longer). Much of the sex-specific mortality reductions were driven by large decreases in pregnancy-related deaths, defined as deaths while pregnant or within two months following pregnancy termination. In part, this may relate to improved engagement in antenatal care and skilled birth attendance[46]. Together with the mortality reduction seen among young children, this suggests that poverty reduction may have had particularly important effects on young families. Indeed, a number of high-profile, government-led cash transfer programmes have focused on pregnant women and young children, either with or without conditions for behaviours such as facility-based delivery[47].

We were not able to differentiate whether people were beneficiaries of a given cash transfer programme because this was not generally elicited in the survey questionnaires, and thus we evaluated changes in mortality for entire populations. Although this might underestimate the effects of cash transfers on direct beneficiaries, our approach has the advantage that it captures spillover effects among non-beneficiaries. For example, cash transfers are often pooled within households, families and even communities[48,49]. Large-scale cash transfer programmes may also affect local and regional economies in favourable ways[25]. This may in part explain why we found population-wide mortality reductions despite many included programmes targeting specific groups (such as older adults or poor families).

This study has several limitations. Because the surveys focus on women of childbearing age, adults over the age of 60 made up only 1% of our adult dataset. Our findings may therefore not apply to these older age groups. Additionally, we were unable to include several populous countries with sizable cash transfer programmes such as Mexico, Brazil and India.

Although we were able to assess heterogeneity across some individual, programme and country factors, the primary contribution of this study remains the overall assessment of the effects of cash transfer programmes across many countries, and the heterogeneity analyses should be considered to be exploratory. In addition, there were other important factors that we were unable to assess that may influence the effectiveness of cash transfer programmes. For example, our study does not address the possibility of implementation quality (programme outreach, enrolment procedures and ‘leakage’ of funds due to corruption, among others) influencing the success or failure of individual programmes. In India, implementation challenges have been cited as a major reason for the failure of some anti-poverty programmes in the past and recent advances in the ability to make secure payments have led to improvements in implementation[50]. The effect of these factors (and other unrelated, granular characteristics) is better assessed through more detailed, programme-specific evaluations, particularly given the lack of comparable implementation data across many countries. Indeed, an important challenge facing many countries is to determine how to improve the design of cash transfer programmes[51,52], including through differing coverage and transfer amounts. For example, recent experimental evidence supports the usefulness of accompanying capital, educational and psychosocial interventions[51]. We did not make cost estimates, so we do not know from our study alone whether the benefits relative to costs of these programmes exceed those of alternative programmes.

Finally, although we attempted to control for confounding through the inclusion of fixed effects and other time-varying covariates, as with any observational study, the possibility of residual confounding remains. Recent advances in the difference-in-differences approach have highlighted instances where findings may be biased, but our use of alternate approaches that are not vulnerable to these biases yielded similar results in our study[38,39,41].

In conclusion, we found that cash transfer programmes were associated with important reductions in the risk of death among adult women and young children across many LMICs. Our findings support the use of such anti-poverty programmes, which many countries have introduced or expanded

during the COVID-19 pandemic, to improve population health and reduce mortality.

Methods

We performed analyses of changes in adult and child mortality associated with implementation of cash transfer programmes between 2000 to 2019, a study period when many countries introduced cash transfer programmes.

Mortality data

To estimate mortality, we generated two individual-level longitudinal datasets—one for adults aged ≥18 years and one for children aged <18 years—using demographic and health surveys (DHS)[36,37,53]. The DHS are conducted in many LMICs about every five years. They use a two-stage cluster sampling design to produce national and sub-national estimates for a variety of indicators that are representative of their target populations[54]. The first stage involves systematic selection of enumeration areas drawn from census files with probability proportional to population size, and the second stage involves a random sampling of households from each enumeration area. Primary respondents were all female household members of reproductive ages (15–49 years). Procedures and questionnaires for DHS have been reviewed and approved by the ICF Institutional Review Board. All analysed data were anonymized. In accordance with standard procedures for secondary data analysis, the University of Pennsylvania Institutional Review Board waived ethical review.

We used surveys that included a maternal mortality module to create the adult dataset. This module collects information from all primary respondents about every sibling born to her biological mother—sex, current vital status, year of death if deceased, current age (or age at death), and for female siblings whether the death was pregnancy-related (death while pregnant or within two months following termination of pregnancy, irrespective of the cause). As there are limited, heterogeneous, and inconsistent data available about other causes of death, we focus on mortality from all causes. Using previously established methodology[36,37,53], we first restructured the dataset such that there was one observation per sibling, and then again such that

each observation corresponded to one person-year from one sibling. Each observation included a binary variable indicating the sibling's survival status during that person-year. We excluded observations from incomplete years (that is, the year of the survey). To minimize recall bias, we excluded observations earlier than ten years before the survey. We excluded person-years during which a sibling was aged <18 years for the purposes of this adult dataset. Of note, because primary respondents in the DHS were 15–49 years of age, older adults were underrepresented.

We created a child dataset from the same set of surveys using the birth history module, which asked female respondents for information about all births—sex, birth date, survival status, and death date. As above, we constructed a longitudinal dataset with observations at the level of the person-year, including an indicator variable for survival and excluding incomplete years and observations earlier than ten years prior to the survey. We excluded person-years during which a child was >17 years old.

We extracted additional data about the primary respondent (sibling in the adult dataset, mother in the child dataset)—age, rural or urban setting, wealth quintile, and schooling attainment (categorized as none, primary, secondary, or greater than secondary). Respondents were classified into wealth quintiles using the DHS Wealth Index, a composite measure of households' cumulative living standard generated using a principal components analysis based on ownership of certain assets, materials used for housing construction, and types of water access and sanitation facilities[55].

Cash transfer programme data

We identified all major, government-led cash transfer programmes within included countries using previously established methods[26]. We manually searched a variety of sources to identify the programmes as well as the years in which they were implemented, the population targeted by the programmes (for example, older adults, families with young children), whether the programmes had behavioural conditionalities, amounts of annual cash transfers, and most recently available number of beneficiaries[56,57,58,59,60]. Data sources included social protection databases from the World Bank, United Nations, and non-governmental organizations, as well as primary

documentation and reporting from individual programmes. We excluded countries with pre-existing cash transfer programmes at the start of the study period.

We calculated the impoverished population coverage for each programme as the most recent estimate of the number of programme beneficiaries divided by the number of individuals in a country with income less than the international poverty line of US$1.90 per day (2011 purchasing power parity). To do this, we divided the most recent estimate of total household beneficiaries by the impoverished population size. If estimates for total beneficiaries were not available, we multiplied direct beneficiaries by the average household size to estimate total beneficiaries[61]. In general, the number of beneficiaries was available during only a limited number of years. Impoverished population sizes were calculated by multiplying the percentages of the populations with income less than the international poverty line (that is, the poverty headcount) prior to programme implementation by the mid-year population from the year of the total beneficiaries estimate[62]. We used the poverty headcount prior to programme implementation because poverty headcount estimates after programme implementation may be decreased by the programmes themselves. For example, if a cash transfer programme began in 2012, we divided the most recent estimate of beneficiaries (numerator) by the poverty headcount in 2012 (denominator) to calculate the impoverished population coverage.

We also calculated the maximum transfer amounts as percentages of GDP per capita in the most recent year the maximum transfer amounts were reported.

Additional country-level data

We obtained additional time-varying covariates for each country and year that are known to be or are likely to be associated with changes in cash transfer programmes and mortality: GDP per capita[62], total health expenditures per capita[62], life expectancies at birth[62], PEPFAR funding budgeted[63], and six Worldwide Governance Indicators from the World Bank that are composite indicators based on 30 data sources: voice and

accountability, political stability and absence of violence, government effectiveness, regulatory quality, rule of law, and control of corruption[62].

Statistical analysis

We used a difference-in-differences approach, a quasi-experimental technique that can be used to estimate causal effects from observational data by comparing the differences in outcomes between intervention and comparison groups during pre-intervention and post-intervention periods, under an assumption of parallel trends (that is, that in the absence of cash transfer programmes, trends in outcomes would be similar in intervention and comparison countries). To do this, we estimated multivariable modified Poisson regression models with the unit of observation being the person-year and a binary outcome variable indicating whether an individual died in a given year[64].

Our primary explanatory variable was a binary variable set to 1 if a cash transfer programme (or combination of programmes) with total impoverished population coverage greater than 5% was active in the respondent’s country during that year. We were prevented from considering coverage as a continuous, time-varying exposure because beneficiary data were available only during a limited number of years for most programmes. We chose 5% based on our prior analyses showing this threshold was associated with improvements in HIV-related outcomes[26], but conducted subgroup analyses (described below) to explore the association with different levels of coverage. We excluded intervention countries that lacked at least two years of mortality data prior to the cash transfer period.

To optimize our comparison country-years, we excluded from our analysis country-years during which cash transfer programmes (or combination of programmes) were implemented with coverage between 2% and 5%. Comparison country-years were therefore defined as those during which there were no active cash transfer programmes, or cash transfer programmes (or combination of programmes) had coverage <2%.

Our effect measure of interest was the risk ratio denoting the association between the cash transfer programme exposure and mortality. In addition to

overall estimates, we also evaluated the temporal relationship between cash transfer programmes and mortality by creating a series of binary indicators for each year before and after the cash transfer period began.

We included in the models country- and individual-level covariates that were likely to confound relationships between cash transfer programmes and mortality. For country-level covariates, we included GDP per capita, PEPFAR funding budgeted, and three Worldwide Governance Indicators: control of corruption, political stability and absence of violence, and voice and accountability. The other three Worldwide Governance Indicators were left out of the models because they displayed substantial multicollinearity with the other covariates as evidenced by variance inflation factors >5. We also considered inclusion of health expenditures per capita, but this variable was not available for all years and adding it to the models minimally impacted the effect estimates.

For individual-level covariates, we included age and rural or urban setting in all analyses. In the child analyses, we also included sex, mother's age, and birth order. We did not include other individual-level variables that were likely to be affected by receipt of cash transfers and/or potentially mediate relationships between cash transfer programmes and mortality (for example, wealth quintile, schooling attainment).

We included country fixed effects to control for time-invariant differences among countries, and year fixed effects to control for secular trends in mortality. We used robust standard errors clustered at the country level to relax the assumption of independently and identically distributed error terms[65,66].

We stratified the adult mortality analysis by sex because of previously identified sex-specific effects of cash transfers[26,43,44,56], and the child mortality analysis by age (<5 years, 5–9 years, 10–17 years) because of highly varying mortality rates by child age[67].

We explored heterogeneity of the effect of cash transfer programmes using subgroup analyses based on the beneficiary, cash transfer programme design, and country factors. For the beneficiary, we considered wealth quintile (of the sibling for the adult analysis and the mother for the child

analysis), age (for the adult analysis, categorized as 18–40, 41–60, and >60 years), educational attainment (of the sibling for the adult analysis, and the mother for the child analysis), rural or urban setting, and cause of death among women (pregnancy-related versus not pregnancy-related). For cash transfer design features, we considered conditionality (unconditional, mixed, or conditional), and four subgroups characterized by most recent impoverished population coverage above or below the median (30%) and maximum annual transfer above or below the median (11% of GDP per capita). For country factors, we considered subgroups characterized by being above or below the median at the start of the cash transfer period for the following: each of the Worldwide Governance Indicators, current annual healthcare expenditures per capita (US$118 purchasing power parity), and life expectancies at birth (62 years). We also stratified by region (sub-Saharan Africa versus outside of sub-Saharan Africa). Finally, we generated country-specific estimates for adult women to allow for informal evaluations of heterogeneity across a range of dimensions.

We also conducted additional sensitivity analyses. First, we assessed the validity of the parallel trends assumption in two ways. We used the previously described temporal analysis to visualize pre-trends, and we estimated regression models using only data prior to the cash transfer period in each country and including an interaction term between an indicator of whether the country was in the intervention group and a linear time trend.

Second, while we used modified Poisson regression models based on conceptual justifications and to be consistent with prior literature assessing changes in mortality using DHS datasets[36,37,53], we assessed for robustness of the results when using logistic and linear models.

Third, recent advances in difference-in-differences analyses with variation in intervention timing have shown that estimates may be biased particularly if there is heterogeneity in intervention effects over time[38,39,63]. When there is effect heterogeneity only in time since the intervention, this concern can be mitigated through use of temporal analysis with dynamic effect estimates (as described above), although there can still be bias present if there are heterogeneous treatment effects over overall calendar time[68]. To address this, we assessed whether a proposed alternative linear estimator not

vulnerable to this bias was consistent with our primary findings[41]. In addition, this bias tends to be influenced by later country-years during the intervention period, so to assess the possible magnitude of this bias we conducted a sensitivity analysis by repeating the primary analysis after excluding country-years after year 5 of the cash transfer programme[69].

Fourth, we assessed whether individual countries might be outliers for key outcomes by assessing whether estimates for women changed substantially after excluding each country individually.

Fifth, we repeated our primary analyses with inclusion of the respondent's wealth quintile and educational attainment.

We did not use statistical methods to predetermine sample size. We performed statistical analyses using SAS V.9.4, R V.3.5.2 using the ggplot2 and forester packages, and STATA V.17 using the did2s package.

Reporting summary

Further information on research design is available in the Nature Portfolio Reporting Summary linked to this article.

Data availability

The analysed data can be requested from the DHS Program website (individual recode datasets from the included countries from https://www.dhsprogram.com/Data/) or are publicly available from the World Bank (GDP per capita, total health expenditures per capita, life expectancies at birth, and Worldwide Governance Indicators datasets from https://data.worldbank.org/data-catalog/) or PEPFAR (PEPFAR Operating Unit Budgets by Financial Classifications FY04-FY20 dataset from https://data.pepfar.gov/financial). The cash transfer programme dataset is available in the Supplementary Information.

References

1. Gortmaker, S. L. Poverty and infant mortality in the United States. *Am. Sociol. Rev.* **44**, 280–297 (1979).

2. Fiscella, K. & Franks, P. Poverty or income inequality as predictor of mortality: longitudinal cohort study. *BMJ* **314**, 1724 (1997).

3. Galea, S., Tracy, M., Hoggatt, K. J., Dimaggio, C. & Karpati, A. Estimated deaths attributable to social factors in the United States. *Am. J. Public Health* **101**, 1456–1465 (2011).

4. Pritchett, L. & Summers, L. H. Wealthier is healthier. *J. Hum. Resour.* **31**, 841–868 (1996).

5. Fritzell, J., Rehnberg, J., Bacchus Hertzman, J. & Blomgren, J. Absolute or relative? A comparative analysis of the relationship between poverty and mortality. *Int. J. Public Health.* **60**, 101–110 (2015).

6. *The State of Social Safety Nets 2018* (World Bank, 2018).

7. Gentilini, U. et al. Social protection and job respones to COVID-19: a real-time review of country measures. *Open Knowledge Repository Beta* https://openknowledge.worldbank.org/handle/10986/37186 (2022).

8. Guanais, F. C. The combined effects of the expansion of primary health care and conditional cash transfers on infant mortality in Brazil, 1998–2010. *Am. J. Public Health.* **105**, S593–S599 (2015).

9. Rasella, D., Aquino, R., Santos, C. A. T., Paes-Sousa, R. & Barreto, M. L. Effect of a conditional cash transfer programme on childhood mortality: a nationwide analysis of Brazilian municipalities. *Lancet* **382**, 57–64 (2013).

10. Shei, A. Brazil's conditional cash transfer program associated with declines in infant mortality rates. *Health Affairs.* **32**, 1274–1281 (2013).

11. Ramos, D. et al. Conditional cash transfer program and child mortality: a cross-sectional analysis nested within the 100 Million Brazilian

Cohort. *PLoS Med.* **18**, e1003509 (2021).

12. Hernández, B., Ramírez, D., Moreno, H. & Laird, N. in *Evaluación Externa del Impacto del Programa Oportunidades* (eds Hernández Prado, B. & Hernández Ávila, M.) 57–76 (2003).

13. Barham, T. & Rowberry, J. Living longer: the effect of the Mexican conditional cash transfer program on elderly mortality. *J. Dev. Econ.* **105**, 226–236 (2013).

14. Rasella, D. et al. Long-term impact of a conditional cash transfer programme on maternal mortality: a nationwide analysis of Brazilian longitudinal data. *BMC Med.* **19**, 127 (2021).

15. Gibson, M. et al. Housing and health inequalities: a synthesis of systematic reviews of interventions aimed at different pathways linking housing and health. *Health Place* **17**, 175–184 (2011).

16. Ivers, L. C. *Food Insecurity and Public Health* (CRC Press, 2015).

17. Ravesteijn, B., van Kippersluis, H. & van Doorslaer, E. The contribution of occupation to health inequality. *Res. Econ. Inequal.* **21**, 311–332 (2013).

18. Feinstein, L., Sabates, R., Anderson, T. M., Sorhaindo, A. & Hammond, C. What are the effects of education on health. In *Proc. of the Copenhagen Symposium* 171–354 (OECD, 2006).

19. Wagstaff, A. Poverty and health sector inequalities. *Bull. World Health Organ.* **80**, 97–105 (2002).

20. Poverty and Inequality Platform. https://pip.worldbank.org/home (World Bank, accessed 15 April 2022).

21. Mahler, D., Yonzan, N., Lakner, C., Castaneda Aguilar, R. & Wu, W. *Updated Estimates of the Impact of Covid-19 on Global Poverty: Turning the Corner on the Pandemic in 2021?* https://blogs.worldbank.org/opendata/updated-estimates-impact-covid-19-global-poverty-turning-corner-pandemic-2021 (World Bank, 2022).

22. Gentilini, U. *Cash Transfers in Pandemic Times: Evidence, Practices, and Implications from the Largest Scale Up in History* (World Bank, 2022).

23. Bastagli, F. et al. *Cash Transfers: What Does the Evidence Say? A Rigorous Review of the Impacts of Cash Transfers and the Role of Design and Implementation Features* (Overseas Development Insitute, 2016).

24. Owusu-Addo, E., Renzaho, A. M. N. & Smith, B. J. The impact of cash transfers on social determinants of health and health inequalities in sub-Saharan Africa: a systematic review. *Health Policy Plan.* **33**, 675–696 (2018).

25. Egger, D., Haushofer, J., Miguel, E., Niehaus, P. & Walker, M. General equilibrium effects of cash transfers: experimental evidence from Kenya. *Econometrica* **90**, 2603–2643 (2022).

26. Richterman, A. & Thirumurthy, H. The effects of cash transfer programmes on HIV-related outcomes in 42 countries from 1996 to 2019. *Nat. Hum. Behav.* **6**, 1362–1371 (2022).

27. Mani, A., Mullainathan, S., Shafir, E. & Zhao, J. Poverty impedes cognitive function. *Science* **341**, 976–980 (2013).

28. Weiser, S. D. et al. In *Food Insecurity and Public Health* (ed. Ivers, L. C.) 23–50 (CRC Press, 2015).

29. Haushofer, J. & Fehr, E. On the psychology of poverty. *Science* **344**, 862–867 (2014).

30. Schilbach, F., Schofield, H. & Mullainathan, S. The psychological lives of the poor. *Am. Econ. Rev.* **106**, 435–440 (2016).

31. Fernald, L. C., Gertler, P. J. & Neufeld, L. M. Role of cash in conditional cash transfer programmes for child health, growth, and development: an analysis of Mexico's Oportunidades. *Lancet* **371**, 828–837 (2008).

32. Walque, D., Fernald, L. C., Gertler, P. J. & Hidrobo, H. In *Disease Control Priorities*, 3rd edn (eds Bundy, D. et al.) Ch. 23 (World Bank, 2017).

33. Barham, T. A healthier start: the effect of conditional cash transfers on neonatal and infant mortality in rural Mexico. *J. Dev. Econ.* **94**, 74–85 (2011).

34. Moncayo, A. L., Granizo, G., Grijalva, M. J. & Rasella, D. Strong effect of Ecuador's conditional cash transfer program on childhood mortality from poverty-related diseases: a nationwide analysis. *BMC Public Health* **19**, 1132 (2019).

35. Lim, S. S. et al. India's Janani Suraksha Yojana, a conditional cash transfer programme to increase births in health facilities: an impact evaluation. *Lancet* **375**, 2009–2023 (2010).

36. Bendavid, E., Holmes, C. B., Bhattacharya, J. & Miller, G. HIV development assistance and adult mortality in Africa. *JAMA* **307**, 2060–2067 (2012).

37. Jakubowski, A., Stearns, S. C., Kruk, M. E., Angeles, G. & Thirumurthy, H. The US President's Malaria Initiative and under-5 child mortality in sub-Saharan Africa: a difference-in-differences analysis. *PLoS Med.* **14**, e1002319 (2017).

38. Goodman-Bacon, A. Difference-in-differences with variation in treatment timing. *J. Econometrics* **225**, 254–277 (2021).

39. Callaway, B. & Sant'Anna, P. H. C. Difference-in-Differences with multiple time periods. *J. Econometrics* **225**, 200–230 (2021).

40. Roth, J., Sant'Anna, P. H., Bilinski, A. & Poe, J. What's trending in difference-in-differences? A synthesis of the recent econometrics literature. Preprint at https://arxiv.org/abs/2201.01194 (2022).

41. Gardner, J. *Two-Stage Differences in Differences* https://jrgcmu.github.io/2sdd_current.pdf (2021).

42. Akresh, R., de Walque, D. & Kazianga, H. *Alternative Cash Transfer Delivery Mechanisms: Impacts on Routine Preventative Health Clinic Visits in Burkina Faso*. National Bureau Of Economic Research Working Paper Series No. 17785 (World Bank, 2012).

43. Bonilla, J. et al. Cash for women's empowerment? A mixed-methods evaluation of the Government of Zambia's Child Grant Program. *World Dev.* **95**, 55–72 (2017).

44. Natali, L., Handa, S., Peterman, A., Seidenfeld, D. & Tembo, G. *Making Money Work: Unconditional Cash Transfers Allow Women to Save and Re-Invest in Rural Zambia*. Innocenti Working Paper No. 2016-02 (UNICEF, 2016).

45. Behrman, J. R. & Parker, S. W. Is health of the aging improved by conditional cash transfer programs? Evidence from Mexico. *Demography.* **50**, 1363–1386 (2013).

46. Glassman, A. et al. Impact of conditional cash transfers on maternal and newborn health. *J. Health Popul. Nutr.* **31**, 48–66 (2013).

47. An analysis of Pradhan Mantri Matru Vandana Yojana. *IAS Parliament* https://www.iasparliament.com/current-affairs/an-analysis-of-pradhan-mantri-matru-vandana-yojana (accessed 15 July 2022).

48. Mosoetsa, S. *Eating from One Pot: The Dynamics of Survival in Poor South African Households* (NYU Press, 2011).

49. Duflo, E. Grandmothers and granddaughters: old-age pensions and intrahousehold allocation in South Africa. *World Bank Econ. Rev.* **17**, 1–25 (2003).

50. Muralidharan, K., Niehaus, P. & Sukhtankar, S. Building state capacity: evidence from biometric smartcards in India. *Am. Econ. Rev.* **106**, 2895–2929 (2016).

51. Bossuroy, T. et al. Tackling psychosocial and capital constraints to alleviate poverty. *Nature.* **605**, 291–297 (2022).

52. Benhassine, N., Devoto, F., Duflo, E., Dupas, P. & Pouliquen, V. Turning a shove into a nudge? A "labeled cash transfer" for education. *Am. Econ. J.* **7**, 86–125 (2015).

53. Obermeyer, Z. et al. Measuring adult mortality using sibling survival: a new analytical method and new results for 44 countries, 1974–2006. *PLoS Med.* **7**, e1000260 (2010).

54. *Demographic and Health Survey Sampling and Household Listing Manual* (ICF International, 2012).

55. Rutstein, S. O. & Johnson, K. *The DHS Wealth Index* (ORC Macro, 2004).

56. Bastagli, F., Hagen-Zanker, J., Harman, L. & Barca V. *Cash Transfers: What Does the Evidence Say?* (Overseas Development Institute, 2016).

57. Social Assistance in Africa Data Platform. https://social-assistance.africa.undp.org/country (United Nations Development Program, accessed 11 January 2021).

58. Non-contributory Social Protection Programs Database. https://dds.cepal.org/bpsnc/cct (United Nations Economic Commission for Latin America and the Caribbean, accessed 11 January 2021).

59. ASPIRE: the Atlas of Social Protection Indicators of Resilience and Equity Program Inventory. https://www.worldbank.org/en/data/datatopics/aspire (World Bank, accessed 11 January 2021).

60. Socialprotection.org Programme Database. https://socialprotection.org/discover/programme (accessed 11 Janary 2021).

61. Household size and composition. https://population.un.org/Household/index.html#/countries/840 (United Nations Department of Economic and Social Affairs Population Division, accessed 13 January 2021).

62. World Development Indicators. https://data.worldbank.org/data-catalog/world-development-indicators (World Bank, accessed 12 January 2021).

63. PEPFAR Panorama Spotlight: financial management. https://data.pepfar.gov/financial (PEPFAR, accessed 11 February 2021).

64. Zou, G. A modified poisson regression approach to prospective studies with binary data. *Am. J. Epidemiol.* **159**, 702–706 (2004).

65. Bertrand, M., Duflo, E. & Mullainathan, S. How much should we trust differences-in-differences estimates? *Q. J. Econ.* **119**, 249–275 (2004).

66. Abadie, A., Athey, S., Imbens, G. W. & Wooldridge, J. *When Should You Adjust Standard Errors for Clustering?* (National Bureau of Economic Research, 2017).

67. Mortality estimates and projections by select age groups. https://population.un.org/wpp/Download/Standard/Mortality/ (United Nations Population Division, accessed 6 September 2022).

68. Sun, L. & Abraham, S. Estimating dynamic treatment effects in event studies with heterogeneous treatment effects. *J. Econometrics.* **225**, 175–199 (2021).

69. Jakiela, P. Simple diagnostics for two-way fixed effects. Preprint at https://arxiv.org/abs/2103.13229 (2021).

Acknowledgements

A.R. was supported by the National Institute of Mental Health of the National Institutes of Health under Award Number K23MH131464.

Author information

Authors and Affiliations

1. Division of Infectious Diseases, Department of Medicine, University of Pennsylvania, Philadelphia, PA, USA

   Aaron Richterman

2. Partners in Health, Mirebalais, Haiti

   Christophe Millien

3. Department of Medical Ethics and Health Policy, University of Pennsylvania, Philadelphia, PA, USA

   Elizabeth F. Bair & Harsha Thirumurthy

4. Partners in Health, Kono, Sierra Leone

   Gregory Jerome

5. Partners in Health, Neno, Malawi

   Jean Christophe Dimitri Suffrin

6. Departments of Economics and Sociology, University of Pennsylvania, Philadelphia, PA, USA

   Jere R. Behrman

7. Population Studies Center, University of Pennsylvania, Philadelphia, PA, USA

   Jere R. Behrman & Harsha Thirumurthy

Contributions

The study was conceptualized by A.R., H.T., C.M., G.J., J.C.D.S. and J.R.B. Methodology design was led by A.R., H.T., E.F.B. and J.R.B., and the data curation and formal analyses were conducted by A.R. under the supervision of H.T. Figures were created by A.R. The first draft of the manuscript was written by A.R. and all authors provided critical inputs into the final draft.

Corresponding author

Correspondence to Aaron Richterman.

Ethics declarations

Competing interests

The authors declare no competing interests.

Peer review

Peer review information

Nature thanks Till Bärnighausen, Davide Rasella and the other, anonymous, reviewer(s) for their contribution to the peer review of this work. Peer reviewer reports are available.

Additional information

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Extended data figures and tables

Extended Data Fig. 1 Country inclusion flow diagram.

Flow Diagram showing selection of intervention (N = 16) and comparison (N = 21) countries during our study period of 2000–2019, and reasons for exclusion (red boxes).

Extended Data Fig. 2 Temporal plots of the effects of cash transfer programs on mortality for children aged 5 to 17 years.

Temporal plots showing the associations between cash transfer programs and mortality as a function of the year of the cash transfer period. Effect estimates are adjusted risk ratios and error bars are 95% confidence intervals. Estimates were generated using multivariable modified Poisson models with country and year fixed effects, country-level covariates (GDP per capita, PEPFAR funding budgeted, and three Worldwide Governance Indicators: Control of Corruption, Political Stability and Absence of Violence, and Voice and Accountability), and individual-level covariates (age and rural/urban setting in all models; sex, age of mother, and birth order in child analyses). We used robust standard errors clustered at the country level. The top panel shows estimates for children aged 5 to 9 years (N = 4,818,370 person-years), the bottom panel shows estimates for children aged 10 to 17 years (N = 4,824,891 person-years).

Extended Data Fig. 3 Heterogeneity analyses for adult males.

Forest plot showing subgroup analyses among adult males (N = 15,249,343 person-years), with fully adjusted risk ratios of mortality with 95% confidence intervals generated using multivariable modified Poisson models with country and year fixed effects, country-level covariates (GDP per capita, PEPFAR funding budgeted, and three Worldwide Governance Indicators: Control of Corruption, Political Stability and Absence of Violence, and Voice and Accountability), and individual-level covariates (age and rural/urban setting). We used robust standard errors clustered at the country level. Effect estimates are adjusted risk ratios and error bars are 95% confidence intervals.

Extended Data Fig. 4 Heterogeneity analyses for children aged <5.

Forest plot showing subgroup analyses among children aged <5 years (N = 6,757,284 person-years), with fully adjusted risk ratios of mortality with 95% confidence intervals generated using multivariable modified Poisson models with country and year fixed effects, country-level covariates (GDP per capita, PEPFAR funding budgeted, and three Worldwide Governance Indicators: Control of Corruption, Political Stability and Absence of Violence, and Voice and Accountability), and individual-level covariates

(age and rural/urban setting). We used robust standard errors clustered at the country level. Effect estimates are adjusted risk ratios and error bars are 95% confidence intervals.

Extended Data Fig. 5 Heterogeneity analyses for children aged 5 to 9.

Forest plot showing subgroup analyses among children aged 5 to 9 years (N = 4,818,370 person-years), with fully adjusted risk ratios of mortality with 95% confidence intervals generated using multivariable modified Poisson models with country and year fixed effects, country-level covariates (GDP per capita, PEPFAR funding budgeted, and three Worldwide Governance Indicators: Control of Corruption, Political Stability and Absence of Violence, and Voice and Accountability), and individual-level covariates (age and rural/urban setting). We used robust standard errors clustered at the country level. Effect estimates are adjusted risk ratios and error bars are 95% confidence intervals.

Extended Data Fig. 6 Heterogeneity analyses for children aged 10 to 17.

Forest plot showing subgroup analyses among children aged 10 to 17 years (N = 4,824,891 person-years), with fully adjusted risk ratios of mortality with 95% confidence intervals generated using multivariable modified Poisson models with country and year fixed effects, country-level covariates (GDP per capita, PEPFAR funding budgeted, and three Worldwide Governance Indicators: Control of Corruption, Political Stability and Absence of Violence, and Voice and Accountability), and individual-level covariates (age and rural/urban setting). We used robust standard errors clustered at the country level. Effect estimates are adjusted risk ratios and error bars are 95% confidence intervals.

Extended Data Fig. 7 Country-specific effects of cash transfer programs on mortality among adult females.

Forest plot showing country-specific effects of cash transfers on mortality among adult females (N = 14,994,934 person-years). Estimates were

generated using multivariable modified Poisson models with country and year fixed effects, country-level covariates (GDP per capita, PEPFAR funding budgeted, and three Worldwide Governance Indicators: Control of Corruption, Political Stability and Absence of Violence, and Voice and Accountability), and individual-level covariates (age and rural/urban setting in all moels; sex, age of mother, and birth order in child analyses). We used robust standard errors clustered at the country level.

Extended Data Table 1 Characteristics of adult mortality dataset
Extended Data Table 2 Characteristics of the child mortality dataset

Supplementary information

Supplementary Information

This file contains Supplementary Tables 1–16 and references.

Reporting Summary

Peer Review File

Rights and permissions

Springer Nature or its licensor (e.g. a society or other partner) holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.

Reprints and Permissions

About this article

Cite this article

Richterman, A., Millien, C., Bair, E.F. *et al.* The effects of cash transfers on adult and child mortality in low- and middle-income countries. *Nature* **618**, 575–582 (2023). https://doi.org/10.1038/s41586-023-06116-2

- Received: 22 September 2022
- Accepted: 21 April 2023
- Published: 31 May 2023
- Issue Date: 15 June 2023
- DOI: https://doi.org/10.1038/s41586-023-06116-2

Share this article

Anyone you share the following link with will be able to read this content:

Get shareable link

Sorry, a shareable link is not currently available for this article.

Copy to clipboard

Provided by the Springer Nature SharedIt content-sharing initiative

Cash-transfer programmes reduce mortality in low- and middle-income countries

Nature Research Briefing 31 May 2023

This article was downloaded by **calibre** from https://www.nature.com/articles/s41586-023-06116-2

- Article
- Open Access
- Published: 31 May 2023

A pan-influenza antibody inhibiting neuraminidase via receptor mimicry

- Corey Momont,
- Ha V. Dang,
- Fabrizia Zatta,
- Kevin Hauser,
- Caihong Wang,
- Julia di Iulio,
- Andrea Minola,
- Nadine Czudnochowski,
- Anna De Marco,
- Kaitlin Branch,
- David Donermeyer,
- Siddhant Vyas,
- Alex Chen,
- Elena Ferri,
- Barbara Guarino,
- Abigail E. Powell,
- Roberto Spreafico,
- Samantha S. Yim,
- Dale R. Balce,
- Istvan Bartha,
- Marcel Meury,
- Tristan I. Croll,
- David M. Belnap,
- Michael A. Schmid,

- William Timothy Schaiff,
- Jessica L. Miller,
- Elisabetta Cameroni,
- Amalio Telenti,
- Herbert W. Virgin,
- Laura E. Rosen,
- Lisa A. Purcell,
- Antonio Lanzavecchia,
- Gyorgy Snell ,
- Davide Corti &
- …
- Matteo Samuele Pizzuto

Show authors

Nature **volume 618**, pages 590–597 (2023)

- 7735 Accesses

- 34 Altmetric

- Metrics details

Abstract

Rapidly evolving influenza A viruses (IAVs) and influenza B viruses (IBVs) are major causes of recurrent lower respiratory tract infections. Current influenza vaccines elicit antibodies predominantly to the highly variable head region of haemagglutinin and their effectiveness is limited by viral drift[1] and suboptimal immune responses[2]. Here we describe a neuraminidase-targeting monoclonal antibody, FNI9, that potently inhibits the enzymatic activity of all group 1 and group 2 IAVs, as well as Victoria/2/87-like, Yamagata/16/88-like and ancestral IBVs. FNI9 broadly neutralizes seasonal IAVs and IBVs, including the immune-evading H3N2 strains bearing an N-glycan at position 245, and shows synergistic activity when combined with anti-haemagglutinin stem-directed antibodies. Structural analysis reveals that D107 in the FNI9 heavy chain

complementarity-determinant region 3 mimics the interaction of the sialic acid carboxyl group with the three highly conserved arginine residues (R118, R292 and R371) of the neuraminidase catalytic site. FNI9 demonstrates potent prophylactic activity against lethal IAV and IBV infections in mice. The unprecedented breadth and potency of the FNI9 monoclonal antibody supports its development for the prevention of influenza illness by seasonal and pandemic viruses.

Main

Seasonal influenza viruses represent major causes of severe respiratory tract infections resulting in 300,000–600,000 deaths per year worldwide[3]. IAVs have a vast host tropism that includes different animal species such as swine, bats and various wild birds, which represent the primary reservoir of these viruses. This wide tropism fosters viral evolution and leads to zoonotic infections as well as pandemics[4]. IBVs predominantly circulate in humans and display lower genetic diversity but can cause severe disease especially in children and high-risk individuals[5].

IAV and IBV infectivity relies on two glycoproteins expressed on the viral envelope that work in concert[6]. Haemagglutinin (HA) binds to the sialic acid (SA) receptor and is pivotal for endosomal membrane fusion, leading to virus uncoating[7]. Neuraminidase (NA) is a receptor-destroying enzyme that cleaves SA, facilitating the release of viral particles from infected cells[8]. NA is also crucial for penetration of the virus through the mucus layer by cleaving decoy receptors that are abundantly present in mucins[9,10]. Owing to their complementary roles, HA and NA activities must be finely balanced to allow viral entry and egress from host cells[11,12].

Despite the lower abundance on the viral envelope than HA[8,13], NA represents a key target for neutralizing antibodies. The protective role of anti-NA antibodies has been documented following the 1968 pandemic in which the H3N2 virus carried an NA antigenically related to the previously circulating H2N2 virus[14]. In addition, more recent studies have demonstrated that serum antibodies with NA inhibition (NAI) activity are an independent correlate of protection against IAV infection[15,16,17]. Besides

blocking the interaction of the enzymatic pocket with SA receptors, NA antibodies can also promote Fc-mediated protection in vivo[18]. Although NA neutralizing antibodies can be induced by natural infection, the anti-NA immunity elicited by current influenza vaccines is weak, mainly due to the low abundance and poor stability of the NA antigen, and to the lack of standardization of the amount of NA in approved vaccines[19].

In view of the antigenic diversity of IAVs and IBVs, the identification of monoclonal antibodies (mAbs) covering both viruses and their variants remains challenging. Several anti-HA stem neutralizing mAbs that recognize the antigenically distant groups 1 and 2 HAs of IAVs have been identified. When tested in a therapeutic setting, these mAbs showed no or only limited efficacy[20,21]. However, one of these mAbs is currently being tested in a phase II prophylactic study[22]. Anti-NA mAbs with broad reactivity may represent a complementary and possibly synergistic approach to anti-HA stem mAbs.

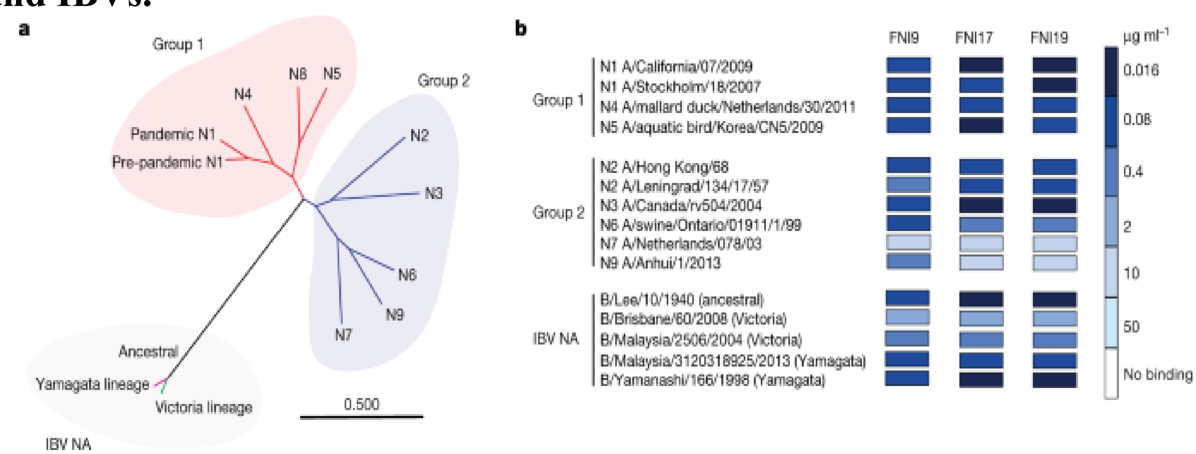
Recent studies described anti-NA neutralizing mAbs isolated following infection or vaccination, showing breadth across several IAVs and IBVs[23,24,25,26]. Here we identify a mAb, designated FNI9, which displays an unprecedented breadth and potency against both IAVs and IBVs via receptor mimicry. These findings together with the synergy with HA stem-directed mAbs support the development of FNI9 for the prevention of influenza illness by seasonal and pandemic viruses.

Isolation of pan-influenza anti-NA mAbs

Sera from almost 200 healthy individuals were screened for the presence of IgG cross-reacting antibodies to non-human circulating group 1 (N4) and group 2 (N3 and N9) NAs transiently expressed on the surface of mammalian cells (Fig. 1a). One donor, designated ‘I’, with high reactivity against these NA antigens was selected. More than 10 million peripheral blood mononuclear cell-derived IgG^+ memory B cells from this donor were interrogated using a N1-based baiting strategy (N1 from H5N1 A/Vietnam/1203/2004) and 14,625 cells were sorted and then seeded as single cells in co-culture with mesenchymal stem cells. Culture supernatants were screened for inhibition of N1 sialidase activity by an enzyme-linked

lectin assay (ELLA). As ELLA also identifies antibodies that inhibit NA activity by steric hindrance, we further screened the NAI activity of the antibodies in the cell supernatants using the 2′-(4-methylumbelliferyl)-α-D-*N*-acetylneuraminic acid (MUNANA) substrate to select mAbs that directly target the enzymatic pocket of group 1 (N1) and group 2 (N2) NAs. Fourteen mAbs (designated FNI mAbs) were identified and found to belong to a single clonal family diversified by a high load of somatic mutations. These antibodies use *IGHV1-69* and *IGK3-15* genes and are characterized by heavy chain complementarity-determinant region 3 (HCDR3) that is 21 amino acids long (Extended Data Fig. 1a–e and Supplementary Table 1).

Fig. 1: Breadth of FNI9, FNI17 and FNI19 anti-NA mAbs across IAVs and IBVs.



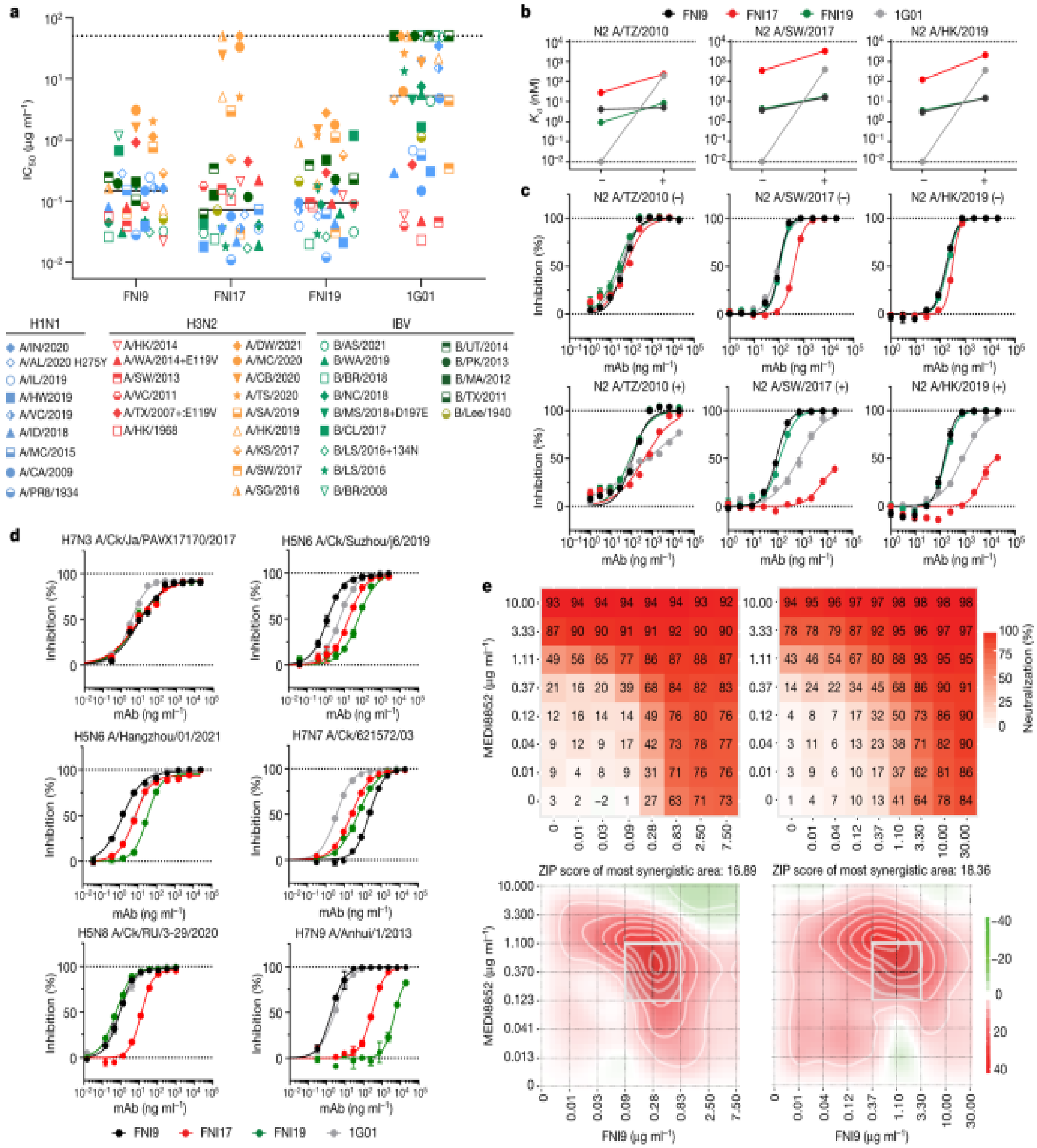
a, Phylogenetic tree of IAV and IBV NAs constructed using maximum likelihood analysis of amino acid sequences. The scale bar indicates the average number of amino acid substitutions per site. **b**, Heat map of mAb binding to cell-surface-expressed NAs representative of IAV and IBV strains by flow cytometry. Coloured boxes represent the lowest concentration at which mAb binding (expressed as mean fluorescence intensity) was measurable.

Three mAbs—FNI9, FNI17 and FNI19—stood out for their binding to a broad panel of cell-surface-displayed seasonal and animal circulating IAV and IBV NAs (Fig. 1a,b, Extended Data Fig. 2a and Supplementary Table 2). To confirm their broad NAI activity, the three mAbs were tested in a MUNANA assay with N1, N2 and IBV NA antigens (Extended Data Fig. 2b). The previously described broadly neutralizing anti-NA mAb 1G01 (ref.

[24]) was also included as a positive control. FNI9 and FNI19 displayed NAI activity similar to 1G01 to N1 and N2 NAs but stronger inhibition of enzymatic activity to the IBV NAs. FNI17 demonstrated the weakest NAI activity versus N2 antigens but the strongest inhibition of IBV NAs (Extended Data Fig. 2b).

We next assessed FNI9, FNI17 and FNI19 mAbs for their neutralizing activity against a panel of seasonal IAV and IBV strains spanning more than 80 years of NA antigenic evolution (Fig. 2a and Supplementary Table 3). The three FNI mAbs efficiently neutralized all IBVs tested, with FNI17 showing the highest potency in line with the NAI results. Conversely, 1G01 mAb neutralized only a fraction of IBVs tested, consistent with its reported limited coverage of IBV strains[24]. Potent neutralization was observed for all the mAbs to seasonal H1N1 strains (Fig. 2a). Of note, 1G01, but not FNI mAbs, was affected by the oseltamivir-resistant mutation H275Y[27] (Fig. 2a), which was present in up to 68% of H1N1 seasonal strains pre-2009 H1N1 pandemic[28]. Although all anti-NA mAbs efficiently neutralized the H3N2 strains isolated before 2015, the activity of FNI17 and 1G01 was strongly impaired by the most recent H3N2 isolates (Fig. 2a). Since the 2014/2015 Northern Hemisphere influenza season, a new N-linked glycosylation site at position 245 was introduced on the N2 NA as a result of the S245N and S247T substitutions (Extended Data Fig. 2c). This glycan, which is located at the entry of the NA enzymatic pocket, has been present since then in all seasonal H3N2 viruses, and has been shown to reduce the activity of antibodies targeting the catalytic site[29,30].

Fig. 2: In vitro characterization of FNI9, FNI17 and FNI19 mAbs.



a, FNI9, FNI17, FNI19 and 1G01 mAbs neutralization half-maximal inhibitory concentration (IC_{50}) values against a panel of seasonal IAVs (H1N1 is blue; H3N2 pre-2015 is in red and H3N2 post-2015 is in orange) and IBVs (B/Victoria/2/87-like viruses are in light green, B/Yamagata/16/88-like viruses are in dark green and ancestral is in olive). Full viral strain designations and IC_{50} values are listed in Supplementary Table 3. Geometric mean of $n = 2$ independent experiments for each strain is

shown. The black solid line indicates the median IC_{50} for each mAb. Symbols on the dotted line represent strains against which the mAbs did not reach IC_{50} at the concentrations tested in the assay (see Supplementary Table 3). **b**, Binding affinity (K_d) of anti-NA Fab fragments to N2 NA antigens not bearing (−) or bearing (+) glycan at position 245 as measured by SPR. Full details of the antigens tested and K_d values are reported in Supplementary Table 4. Results represent an average of at least two technical replicates from one independent experiment. Dotted lines represent the limits of detection. **c**, Inhibition of enzymatic activity, as measured by ELLA, exerted by anti-NA mAbs on NA antigens not bearing (−) or bearing (+) N245 glycan. Full details of the antigens used are reported in Supplementary Table 4. **d**, Inhibition of enzymatic activity, as measured by ELLA, exerted by anti-NA mAbs against NA-only based pseudoparticles bearing N3, N6, N7, N8 or N9 representative of zoonotic isolates. Dotted lines in **c** and **d** represent the minimum and maximum percentage of inhibition. Results represent two technical replicates from one independent experiment out of two. Error bars indicate s.d. of technical duplicates. **e**, In vitro neutralization matrixes (top panels) and synergy plots (bottom panels) reporting combination activity of anti-HA stem-directed MEDI8852 and anti-NA FNI9 mAbs to H1N1 A/Puerto Rico/8/34 (left) and H3N2 A/Tasmania/503/2020 (right) viruses. Neutralization matrixes were performed in technical triplicates with one of two independent experiments shown.

Source Data

To dissect the effect of the N245 glycan on antibody activity, we produced a panel of three N2 antigens derived from H3N2 strains isolated before and after 2015 together with their counterparts in which the glycosylation site was either introduced or removed (Supplementary Table 4). Surface plasmon resonance (SPR) analysis revealed that FNI17 Fab has weaker affinity than FNI9, FNI19 and 1G01 towards N2 antigens even when the glycan N245 was not present, and glycosylation further decreased its binding (Fig. 2b and Supplementary Table 4). Although the 1G01 Fab displayed the highest binding affinity in the absence of the glycosylation, the presence of the N245 glycan severely dampened 1G01 binding to all N2 antigens tested (Fig. 2b). Despite the N245 glycan being associated with decreased binding to N2,

FNI9 and FNI19 retained nanomolar binding affinity (Fig. 2b). Consistent with SPR results, all mAbs showed similar inhibition of N2 enzymatic activity when the N245 glycan was not present, but the introduction of the glycosylation resulted in a marked decrease of NAI activity by FNI17 and 1G01 compared to FNI9 and FNI19 (Fig. 2c). Our results for 1G01 are in line with those recently reported[31], indicating a decrease in NAI activity by this mAb to N2 bearing the N245 glycan, albeit 1G01 in vivo protection was retained at the single dose tested. Overall, FNI9 and FNI19 mAbs were found to be the most active to contemporary H3N2 strains carrying the N245 glycosylation as well as to oseltamivir-resistant strains.

Next, we assessed the ability of the FNI mAbs to inhibit the enzymatic activity of pseudoparticles bearing NAs representative of enzootic and zoonotic avian and mammalian IAVs. FNI mAbs displayed different inhibitory activity against pseudoparticles bearing NAs from highly pathogenic avian IAVs that were previously reported to have infected humans[32] or were responsible of the recent outbreak in farmed mammals[33] (Fig. 2d and Extended Data Fig. 3a). Of note, FNI9 showed the highest NAI activity against N9 from H7N9 among the clonally related antibodies. FNI mAbs displayed comparable NAI activity to pseudoparticles decorated by low pathogenic avian influenza viruses N3, N4 or N5, with the latter not being sensitive to the 1G01 mAb (Extended Data Fig. 3b). Finally, the FNI9 mAb demonstrated broad inhibitory activity across a panel of pseudoparticles presenting NAs from enzootic IAVs circulating in swine and dogs (Extended Data Fig. 3c).

Besides inhibiting enzymatic activity and viral replication in vitro, FNI9 induced complement-dependent cell cytotoxicity, mediated the lysis of infected cells by human primary natural killer cells (antibody-dependent cell cytotoxicity) and promoted antibody-dependent cell phagocytosis by monocytes (Extended Data Fig. 4).

The potential interaction of anti-NA mAbs with other influenza virus-targeting antibodies was investigated by combining FNI9 with the broadly neutralizing HA-stem directed mAb MEDI8852 (ref. [34]). In a checkerboard neutralization assay, these mAbs were found to strongly synergize with a 1:1 anti-HA:anti-NA ratio against both H1N1 A/Puerto Rico/8/34 and H3N2 A/Tasmania/503/2020 (Fig. 2e), suggesting that synergistic activity of

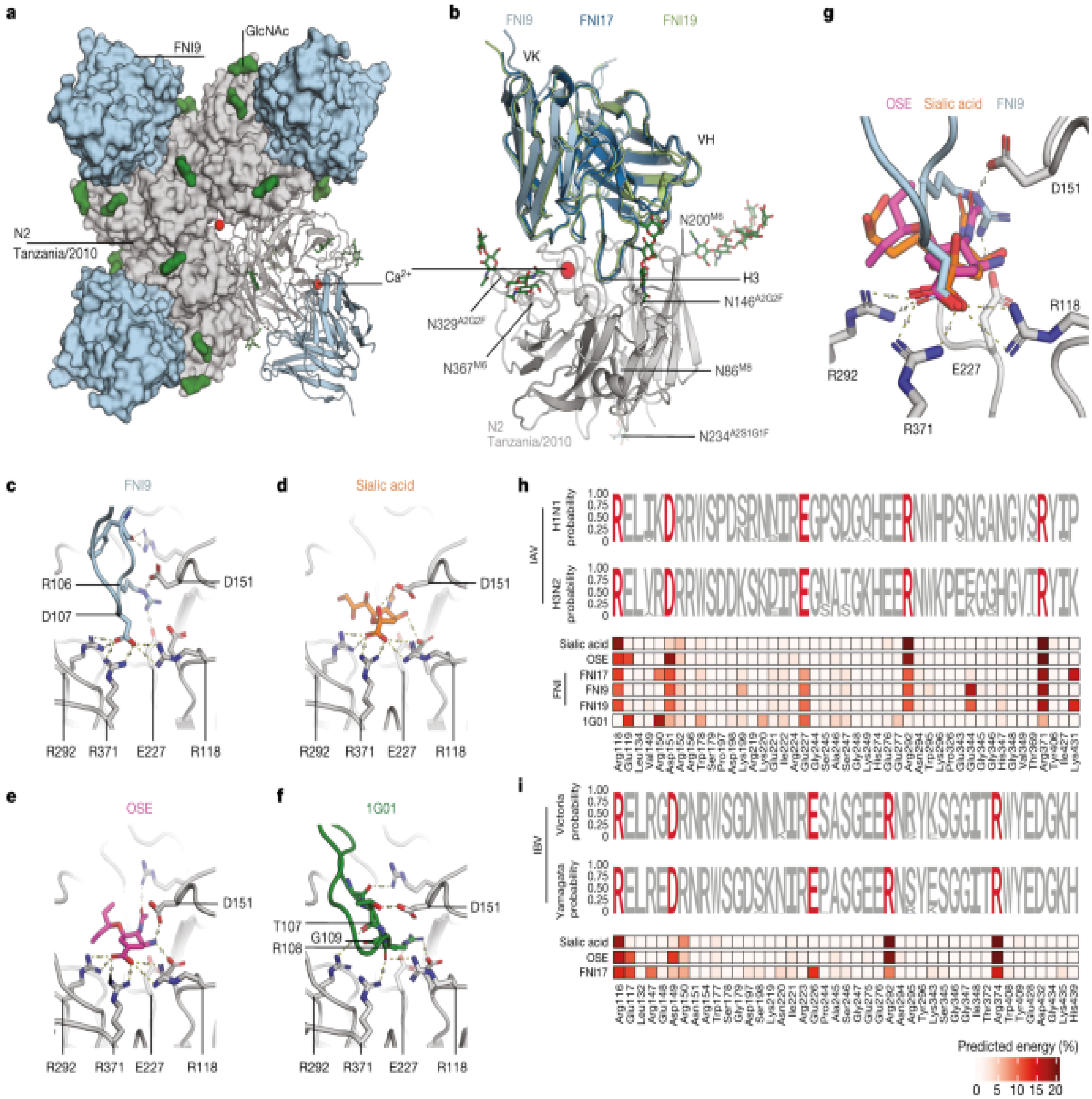
broadly neutralizing mAbs could be exploited to increase the breadth and barrier to viral resistance.

Collectively, we identified broadly neutralizing NA-targeting mAbs with unprecedented breadth to seasonal and zoonotic IAVs and IBVs, which can synergize with anti-HA stem-directed mAbs.

Receptor mimicry by FNI mAbs

To elucidate the structural basis for the FNI9, FNI17 and FNI19 broad neutralizing activity, we determined the cryo-electron microscopy structures of their Fabs in complex with NA from H3N2 A/Tanzania/205/2010 at 2.9 Å, 2.3 Å and 3.1 Å resolutions, respectively (Supplementary Table 5). The overall topology of the FNI9–NA complex adopts a C4 symmetry, with each NA protomer bound to a Fab (Fig. 3a). FNI9, FNI17 and FNI19 all adopt the same binding mode (Fig. 3b), inserting the HCDR3 into the enzymatic pocket where residues R106 and D107 establish multivalent salt bridges with the NA active site residues R118, D151, R292 and R371 (ref. [35]) (Fig. 3c).

Fig. 3: Structure of anti-NA mAbs targeting the SA-binding site.



a, Topology of the complex formed by FNI9 (light blue) binding to NA (grey). Representative calcium ions at the centre of the tetramer and in the NA–Fab interface are represented as red spheres. Glycans decorating each NA protomer are shown in green (see Extended Data Fig. 7). **b**, Binding of FNI9 (light blue), FNI17 (dark blue) and FNI19 (light green) to the NA (grey) SA-binding pocket. Only the variable domains of the mAbs are shown. VH, variable domain heavy chain; VK, variable domain kappa light chain. **c**, Network of salt bridges and hydrogen bond interactions between R106 and D107 (light blue sticks) in the HCDR3 of FNI9 and R118, D151, E227, R292 and R371 in NA (grey sticks). The dashed lines depict interactions within 3.5 Å. **d**, Network of salt bridges and hydrogen bonds

between SA (orange) and NA-binding pocket residues (grey) based on PDB: 4GZQ. The dashed lines are depicted as in **c**. **e**, Network of salt bridges and hydrogen bonds between oseltamivir (OSE; pink) and NA-binding pocket residues (grey) based on PDB: 4GZP. The dashed lines are depicted as in **c**. **f**, Hydrogen bonds between residues T107, R108 and G109 (green) in the HCDR3 of 1G01 (PDB: 6Q23) and the NA-binding pocket residues (grey). The dashed lines are depicted as in **c**. **g**, Overlay of SA, OSE and the HCDR3 of FNI9 illustrates the molecular mimicry of their carboxylates participating in a tridentate salt bridge with R118, R292 and R371. The dashed lines are depicted as in **c**. **h**,**i**, Logo plot amino acid conservation of SA, OSE, FNI9, FNI17, FNI19 and 1G01 epitopes based on available NA sequences from human seasonal H1N1 ($n = 64{,}476$) and H3N2 ($n = 91{,}754$) IAVs (**h**) and Victoria/2/87-like ($n = 23{,}787$) and Yamagata/16/88-like ($n = 17{,}769$) IBVs (**i**). Key contact residues are shown in red. The binding energies of epitope residues (as a percentage of the total) are shown in the heat map; for the IBV analyses, PDB 1NSC (SA, B/Beijing/1/87) and 4CPY (OSE, B/Lyon/CHU/15.216/2011) were used.

Source Data

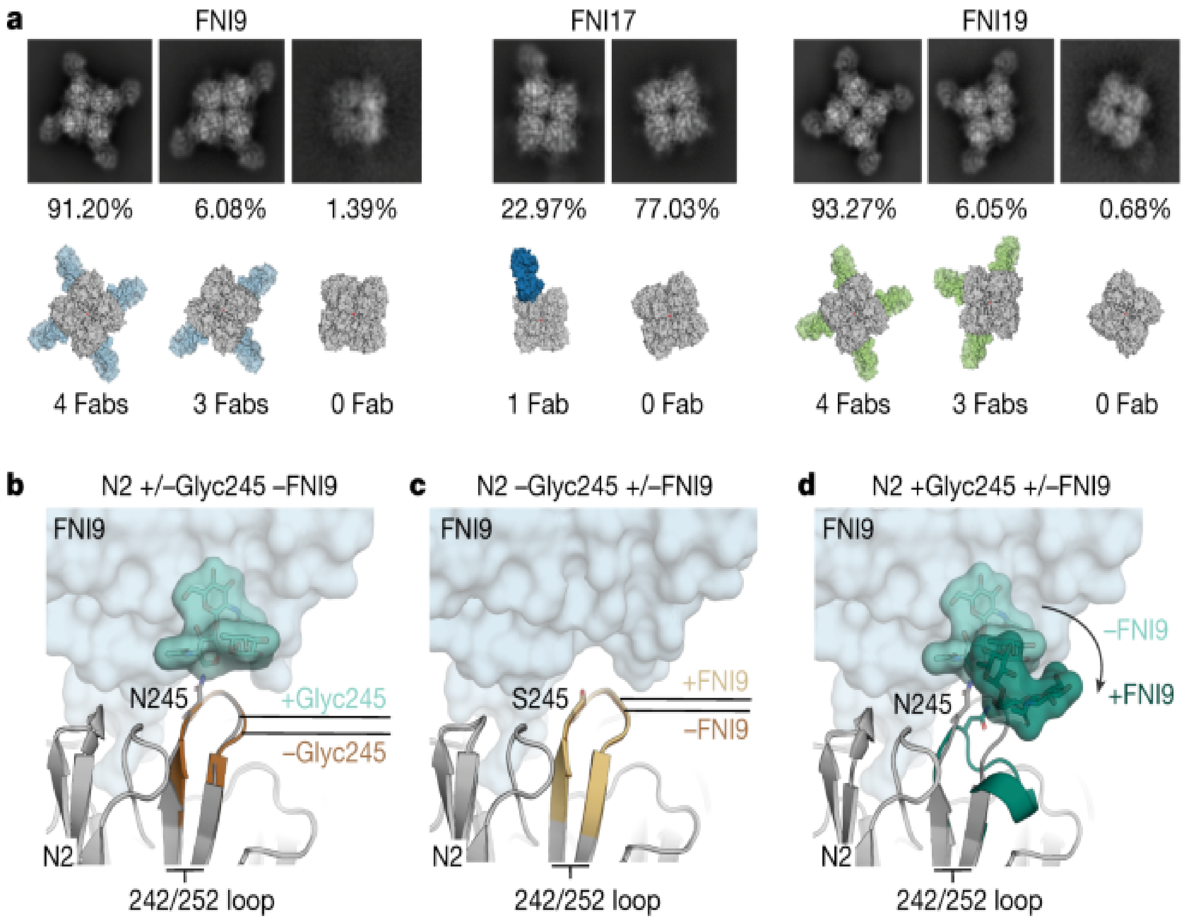
These four NA active site residues also form the primary interactions with SA, suggesting that the broad neutralizing activity of FNI mAbs is the result of receptor molecular mimicry (Fig. 3d,g), similar to the binding mode of oseltamivir (Fig. 3e,g). By contrast, 1G01 interactions within the enzymatic pocket are distinct from those observed for the SA receptor, oseltamivir and the FNI mAbs (Fig. 3f). Of note, both 1G01 and FNI mAbs also interact with E227, which contributes to form the NA active site framework (Fig. 3c,f). Owing to the functional constraints associated with the SA receptor interactions, the key contact residues of the FNI mAbs are highly conserved across IAV and IBV NAs (Fig. 3h,i and Extended Data Fig. 5). Moreover, in vitro resistance studies with live virus as well as deep mutational scanning via an NA library expressed in mammalian cells revealed a very limited number of NA mutations capable of reducing, but not abrogating, the binding or activity of FNI mAbs (Extended Data Fig. 6 and Supplementary Table 4). Of note, most of these mutations were only retrieved after several rounds of viral passages, are associated with a decrease in the NA sialidase activity and are extremely rare in human and animal isolates, suggesting a

fitness cost for the corresponding virus mutants and a high barrier to resistance for this class of anti-NA antibodies (Extended Data Fig. 6).

To obtain a more detailed understanding of the binding epitopes of FNI mAbs, molecular dynamics (MD) simulations of FNI–NA complexes were performed using the complete glycans revealed by peptide mapping liquid chromatography–mass spectrometry (Extended Data Fig. 7). Although the analysis of the static structures shows similar binding energies for the FNI9–NA and FNI17–NA complexes (Extended Data Fig. 8a), the MD-derived total dynamic binding energy predicts that FNI9 is a stronger binder (−370 MOE (molecular operating environment) kcal per mol) than FNI17 (−250 MOE kcal per mol) for NA from the H3N2 A/Tanzania/205/2010 strain (Extended Data Fig. 8b,c), in line with binding affinity measurements (Fig. 2b and Supplementary Table 4). The interactions are balanced across more residues in the FNI9–NA complex than in the FNI17–NA complex both in the epitope (Extended Data Fig. 8d) and in the paratope (Extended Data Fig. 8e), providing a possible mechanistic explanation for the superior binding of FNI9.

We next investigated the structural basis for the reduced binding affinities to N2 antigens bearing the N245 glycan. Two-dimensional classification of the cryo-electron microscopy data showed the presence of N2 tetramers carrying the N245 glycan saturated with three or four FNI9 Fabs (6.08% with three FNI9 Fabs and 91.20% with four FNI9 Fabs) or FNI19 Fabs (6.05% with three FNI19 Fabs and 93.27% with four FNI19 Fabs), whereas only zero or one FNI17 Fabs (77.03% with zero FNI17 Fabs and 22.97% with one FNI17 Fab) were found to occupy each tetramer (Fig. 4a). These data are consistent with the lower binding affinity and neutralizing activity of FNI17 than FNI9 and FNI19 versus the N2 antigens and viruses bearing the N245 glycan (Fig. 2a,b).

Fig. 4: FNI mAbs induce a conformational change in the 242/252 loop when N245 is glycosylated.



a, 2D and 3D classifications and percentages of classes of FNI9–NA (N2 A/Hong Kong/2019) showing four, three and zero Fabs bound to the tetramer (the condition with a single Fab bound is not shown: 1.33%), FNI17–NA (N2 A/Tanzania/2010 with S245N and S247T) showing one and zero Fabs bound to the NA tetramer, and FNI19–NA (N2 A/Hong Kong/2019) with zero, three and four Fabs bound to the NA tetramer. **b**, N245 glycosylated (+Glyc245; grey) and non-glycosylated (−Glyc245; brown) NAs reveal very similar conformations of the 242/252 loop in the unbound state. Overlay of the FNI9 Fab (light blue, translucent surface) indicates the steric hindrance with the N245 glycan (light green surface), which would occur without a rearrangement of the 242/252 loop and the glycan. **c**, Overlays of N2 A/Tanzania/2010 NA (−Glyc245) with (+FNI9, gold) and without (−FNI9, brown) Fab bound reveal that the NAs adopt indistinguishable conformations; the FNI9 Fab is shown as a light blue surface. **d**, Overlays of N245-glycosylated NA structures with (+FNI9, dark green) and without (−FNI9, grey) Fab bound illustrate the influence of the Fab on inducing a conformational change in the 242/252 loop. Although the

peptide mapping liquid chromatography–mass spectrometry in Extended Data Fig. 7 revealed that position 245 bears an A2G2F glycan, only the two GlcNAcs and fucose resolved by cryo-electron microscopy are shown in the figure.

Source Data

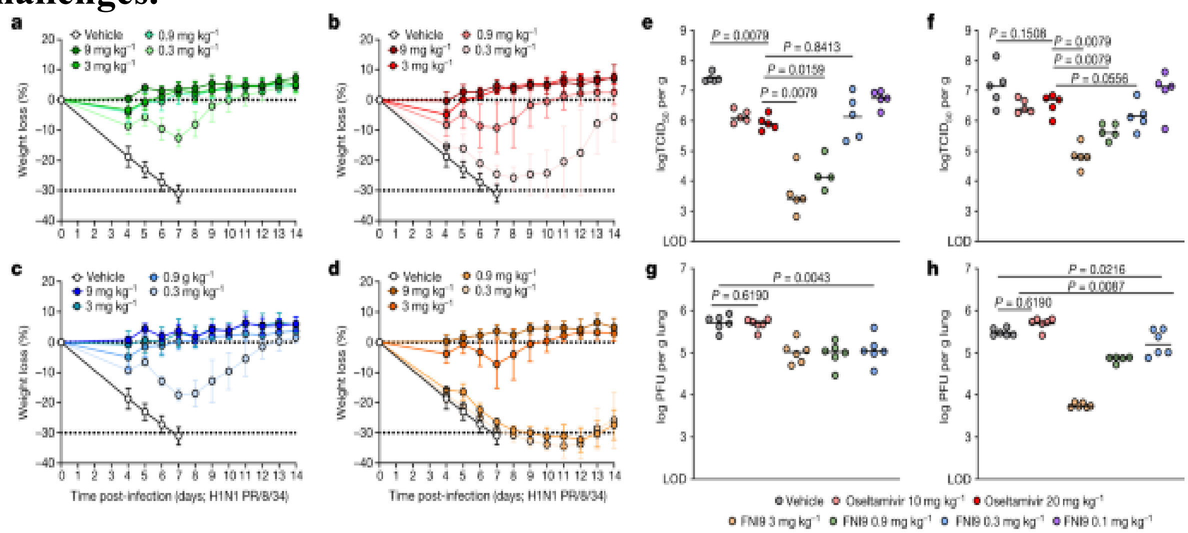
Structural analysis revealed that the conformation of the 242/252 loop, which contains the $N_{245}AT_{247}$ glycosylation motif, is not influenced by the presence or absence of the N245 glycan (Fig. 4b). In N2 NA structures without the glycan, there is no clash with the antibody and the 242/252 loop adopts virtually indistinguishable conformations in both bound and unbound states (Fig. 4c). By contrast, following binding of the FNI9 antibody, both the 242/252 loop and the N245 glycan undergo a dramatic conformational change to avoid a clash between the Fab and the glycan (Fig. 4d). The energy required for the induced fit of loop 242/252 could underlie the differences in affinity and neutralizing activity of the FNI mAbs against the recent N2 NAs.

Prophylactic activity of the FNI9 mAb

To assess the in vivo activity of FNI9 in a mouse model, we produced a murinized IgG2a form of this mAb (muFNI9) and a N297Q mutant (muFNI9(N297Q)) that lacks FcRs and complement C1q binding[36]. For comparison, we also produced the same murinized versions of MEDI8852, a broadly IAV neutralizing anti-HA stem antibody[34]. The four mAbs were prophylactically administered via intravenous injection to BALB/c mice 24 h before lethal challenge with H1N1 A/Puerto Rico/8/34. Mice receiving muFNI9 were completely protected from loss of body weight down to 0.9 mg kg^{-1} and displayed limited loss of body weight at 0.3 mg kg^{-1}, with all the animals recovering throughout 14 days (Fig. 5a and Extended Data Fig. 9a). Administration of muMEDI8852 conferred full protection from body weight loss down to 3 mg kg^{-1}; whereas moderate-to-significant body weight loss was observed at 0.9 and 0.3 mg kg^{-1}, respectively (Fig. 5b and Extended Data Fig. 9a). When the same antibodies were tested with the N297Q mutation, muFNI9(N297Q) displayed protection comparable with that of the parental antibody at all the doses tested (Fig. 5c and Extended

Data Fig. 9a). By contrast, muMEDI8852 bearing N297Q showed a marked decreased activity at limiting doses (0.9 and 0.3 mg kg^{-1}), consistent with previous findings on the role of effector function in protection by anti-HA stem-directed antibodies[37] (Fig. 5d and Extended Data Fig. 9a). Overall, these data indicate that FNI9 has strong prophylactic activity against lethal infection with H1N1 A/Puerto Rico/8/34, which does not appear to be dependent on Fc-mediated effector functions at the doses tested.

Fig. 5: FNI9 mAb protects mice from seasonal IAV and IBV lethal challenges.



a–**d**, Percentage of body weight loss of BALB/c mice (*n* = 6 mice per group) prophylactically administered with murinized anti-NA FNI9 mAb (**a**), anti-HA stem-directed MEDI8852 mAb (**b**) and the same mAbs bearing the N297Q Fc mutation (**c**,**d**) 24 h before lethal infection with H1N1 A/Puerto Rico/8/34. Doses are reported in different colours and the average body weight loss for each dose group is shown. Error bars represent standard deviations. The 0% dotted line indicates baseline body weight loss; the −30% dotted line indicates body weight loss % for euthanasia based on lead veterinarian assessment as described in the ethical statement. **e**–**h**, Replicating virus titres in the lungs of BALB/c mice measured 4 days post-infection with H1N1 A/Puerto Rico/8/34 (**e**), H3N2 A/Singapore/INFIMH-16-0019/2016 (**f**) (*n* = 5 mice per group), B/Victoria/504/2000 (Yamagato linage) (**g**) and B/Brisbane/60/2008 (Victoria lineage) (**h**) (*n* = 6 mice per group) following prophylactic administration of FNI9 at 3, 0.9, 0.3 mg kg^{-1} (**e**–**h**) and 0.1 mg kg^{-1} (**e**,**f**). Each panel presents data derived from *n* = 1

independent experiment. Two-tailed Mann–Whitney test was used for statistical analysis of significance. LOD, limit of detection; TCID, tissue culture infectious dose.

Source Data

Next, we observed that mice prophylactically administered with an ineffective dose of MEDI8852 combined with FNI mAbs in a 1:1 ratio were overall better protected from morbidity than those receiving only the same amount of FNI mAbs contained in the mix, suggesting that the activity of anti-HA stem antibodies can benefit from the association with NA-inhibiting antibodies (Extended Data Fig. 9b).

Given the similarity in the mechanism of action, we compared the prophylactic activity of a single injection of FNI9 with daily administration of oseltamivir in BALB/c mice lethally challenged with a panel of IAVs and IBVs (Fig. 5e–h). Daily administration of oseltamivir at 10 mg kg^{-1} (ref. [38]) reduced viral titre by approximately 1 log against IAVs (H1N1 A/Puerto Rico/8/34 or H3N2 A/Singapore/INFIMH-16-0019/2016) and no further viral titre reduction was observed by doubling the dose (20 mg kg^{-1}) of the small molecule. Conversely, FNI9 at 3 and 0.9 mg kg^{-1} reduced lung viral titres by 4 and 3 logs after challenge with H1N1 A/Puerto Rico/8/34 and by 3 and 2 logs after challenge with H3N2 A/Singapore/INFIMH-16-0019/2016 that carries the N245 glycan (Fig. 5e,f). A viral lung titre reduction comparable with the one provided by oseltamivir was obtained with both IAVs when FNI9 was administered at 0.3 mg kg^{-1} (Fig. 5e,f).

In line with the lower in vitro potency and clinical efficacy against IBVs than IAVs[39,40], oseltamivir did not show significant activity to IBVs B/Victoria/504/2000 (B/Yamagata/16/88-like virus) and B/Brisbane/60/2008 (B/Victoria/2/87-like virus). Conversely, administration of FNI9 significantly reduced viral lung titres to B/Victoria/504/2000 and B/Brisbane/60/2008 down to 0.3 mg kg^{-1} (Fig. 5g,h). Of note, viral replication in the lungs inversely correlated with the serum mAb concentration measured 2 h before infection (Extended Data Fig. 10).

Collectively, these results indicate that a single administration of FNI9 mAb at 0.3 mg kg^{-1} is highly efficacious in preventing replication of both IAVs and IBVs in the lungs of BALB/c mice.

Discussion

The effectiveness of current influenza vaccines is often suboptimal due to mismatch with circulating seasonal strains, immunological imprinting and limited immunogenicity, especially in elderly and immunocompromised individuals[2,41]. Previous studies have shown that NA-directed antiviral drugs are effective in prophylaxis[42,43], albeit their use in this setting is limited by their short half-life and potential selection of escape mutants. Furthermore, anti-NA serum antibodies have been shown to independently correlate with protection from IAV infection[15,44]. In comparison with other previously described anti-NA mAbs[23,24], the FNI9 mAb presented in this study displays larger breadth and higher potency across seasonal IAVs and IBVs, including contemporary H3N2 strains. The SA receptor molecular mimicry exploited by the FNI9 mAb might provide an additional layer of protection from viral escapes, as the key interacting residues in the epitope are evolutionarily conserved in human and animal IAVs and IBVs and most of the NA mutations retrieved from in vitro resistance studies are associated with a fitness cost for the virus. Of note, molecular mimicry of the SA receptor was also observed for anti-HA mAbs[45], which displayed limited breadth. The features of the FNI9 mAb combined with half-life extension through Fc modifications[46] support its development as a single dose per season prophylaxis against influenza.

The synergistic activity observed when FNI9 is combined with the MEDI8852 mAb is consistent with previous reports on the synergy between oseltamivir and anti-HA stem mAbs[34]. In addition, other studies have shown that anti-NA mAbs can increase the Fc-mediated protection by anti-HA stem mAbs[47,48]. We therefore speculate that FNI9 efficacy could be further increased by synergy with endogenous anti-HA stem antibodies elicited by previous infections or vaccinations.

Collectively our findings support the notion that NA is an important target for influenza prophylaxis and justify ongoing efforts to produce recombinant stabilized NA immunogens[49,50] to elicit broadly reactive NAI antibodies, thus contributing to the development of a universal influenza vaccine.

Methods

Cells and viruses

Cell lines were obtained from the American Type Culture Collection (ATCC) (HEK293T/17, Madin-Darby Canine Kidney (MDCK), MDCK-London (MDCK-LN) and A549), or Thermo Fisher Scientific (ExpiCHO-S and Expi293F cells). Expi293F and ExpiCHO-S cells were maintained in Expi293 Expression Medium (Gibco) and ExpiCHO Expression Medium (Gibco), respectively. FreeStyle 293-F cells (Thermo Fisher) used in the DMS experiments were cultured in pre-transduction FreeStyle™ 293 Expression Medium (Thermo Fisher, 12338018) or post-transduction media consisting of DMEM + 10% FBS +1% HEPES + 1% Pen/Strep, with 1.5 μg/mL puromycin for selection. All cell lines used in this study were routinely tested for mycoplasma and found to be mycoplasma-free. Wild-type influenza strains were obtained from the International Reagent Resource (CDC) and the Centre for Biological Reference Materials (NIBSC). Propagation of viral stocks was performed as previously described[51]. In brief, MDCK-LN cells (FR-58, International Reagent Resource) were seeded at 6×10^6 cells in a T75 flask in growth medium (DMEM (11995040, Gibco), 10% FBS (97068-085, VWR), 0.01 M HEPES (15630-080, Gibco), 100 U ml^{-1} penicillin–100 μg ml^{-1} streptomycin (15140-122, Gibco)) and infected the next day at a multiplicity of infection (MOI) of 0.1 in infection medium. Cells were washed twice with PBS before infection. Virus was diluted in infection medium (DMEM, 0.1% BSA, 0.01 M HEPES, 100 U ml^{-1} penicillin–100 μg ml^{-1} streptomycin) and added to cells for 1 h at 37 °C. After absorption, virus inoculum was removed, cells were washed once with PBS, and 10 ml of infection medium containing 2 μg ml^{-1} of l-1-tosylamido-2-phenylethyl chloromethyl ketone (TPCK)-treated trypsin (100 mg; LS003740, Worthington Biochemical) was added to the flasks. The infected cells were incubated at 37 °C for IAV and 35 °C for

IBV in a tissue culture incubator for 24 h. Infection medium were then harvested, spun at 2,000*g* for 5 min, and supernatant was collected, aliquoted and stored at −80 °C. Virus titres were determined by focus forming units (FFU) per millilitre, as described below.

Sample donors

Samples from healthy individuals were obtained from Swiss blood donation centres. Blood donors agreed, by means of a written inform consent form, that their blood or certain components of it could be used anonymously for research purposes. Anonymous blood samples were banked and used for research at VIR Bellinzona in the context of study protocols approved by the local Institutional Review Boards (Canton Ticino Ethics Committee, Switzerland).

B cell isolation and stimulation

Memory B cells from the selected donors were isolated from cryopreserved peripheral blood mononuclear cells (PBMCs) as follow: $CD19^+$ B cells were enriched from PBMCs by staining with phycoerythrin (PE)-Cy7-labelled anti-human CD19 (BD Biosciences) on ice for 30 min, followed by staining with anti-PE beads (Miltenyi Biotec) on ice for 20 min and, finally, positive selected on MACS-separation LS columns (Miltenyi Biotec). NA1 from IAV (N1-IAV) reactive memory IgG^+ B cells were isolated by staining with a biotinylated recombinant N1 (H5N1 IAV-NA; VIVA Biotech) in complex with streptavidin-AF647 (Life Technologies), anti-human IgA PE (Southern Biotech), anti-human IgM PE (BioLegend) and anti-human IgD PE (BD Bioscience) on ice for 45 min and by fluorescence-activated cell sorting (FACS) on a FACSAria (BD Biosciences). Cells were resuspended in IMDM with 10% FBS in the presence of 2.5 μg ml^{-1} CpG 2006, 500 U ml^{-1} IL-2, 5 ng ml^{-1} IL-6 (BD Pharmingen), 50 ng ml^{-1} IL-10 (ImmunoTools) and 10 ng ml^{-1} IL-21 (ImmunoTools) and seeded under clonal conditions on a monolayer of mesenchymal stem cells in 384-well microtitre plates (Corning). After 8–10 days, the supernatants were screened for their NAI activity in a high-throughput ELLA. The B cell supernatants that were having NAI activity were re-tested in a secondary screening against N1, N2 and N9 (VIVA Biotech) in a NAI assay MUNANA (NI-MUNANA). Cross-

reactive memory B cells were lysed, and the variable regions of the heavy and light chain were cloned.

Sequence analysis of mAbs

Complementary DNA (cDNA) was synthesized from selected B cell culture and both the heavy and the light chain variable region (VH and VL) sequences were retrieved by PCR as previously described[52]. Using the database IMGT (http://www.imgt.org), the VH and VL gene family and the number of somatic mutations were determined by analysing the homology of the VH and VL sequences to known human V, D and J genes. Unmutated common ancestor (UCA) sequences of the VH and VL were constructed using IMGT/V-QUEST.

The DNA sequences of the VH and VL regions of 14 clonally related mAbs (FNI1–FNI20) were aligned using the online multiple sequence alignment program Clustal Omega (https://www.ebi.ac.uk/Tools/msa/clustalo/). Subsequently, the UCA sequence was generated using the UA inference application, and the generation and analysis of the phylogenetic tree were performed with dnamI software and AncesTree software[53], respectively. A phylogenetic tree of the concatenated heavy and light chain sequences was generated assuming the same single somatic substitution process across both genes. Each heavy chain sequence was appended to the light chain sequence of the same clone for the purpose of phylogenetic reconstruction under the assumption of a single substitution process. A phylogenetic tree and ancestral states were inferred using raxml-ng with the Blosum62 amino acid substitution model[54].

Production of recombinant mAbs

The VH and VL regions were cloned into IgG1 and IgK expression vectors[52] and expressed recombinantly by transient transfection of ExpiCHO cells (Thermo Fisher Scientific) using the ExpiFectamine CHO transfection kit (Thermo Fisher Scientific). After 8 days, cell culture supernatants were purified by protein A chromatography on ÄKTA Xpress Fast Protein Liquid Chromatography (Cytiva) using HiTrap Protein A columns (Cytiva) followed by buffer exchange to histidine buffer using HiPrep 26/10 desalting

columns (Cytiva). The purified antibodies were quantified using the MAbPac protein A column (Thermo Fisher Scientific). The FNI17 and FNI19 Fabs were recombinantly produced by transient transfection of ExpiCHO using the same conditions reported for the full IgG antibodies. The FNI9 Fab was produced in HEK293 suspension cells. Purification was performed using CaptureSelect IgG-CH1 resin (5943462005, Thermo Scientific) followed by buffer exchange into PBS. The 1G01 Fab was obtained by fragmentation of the corresponding IgGs using the FabALACTICA Fab kit (A2-AFK-025, Genovis) according to the manufacturer's protocol.

Transient expression of NA proteins on Expi cells

IAV and IBV NA protein coding sequences were cloned into the phCMV1 expression plasmid (Genlantis) under the control of the human CMV promoter. Immediately before transfection, ExpiCHO-S or Expi293F cells were seeded at 6×10^6 cells per millilitre or 3×10^6 cells per millilitre, respectively, in a volume of 5 ml in a 50-ml bioreactor. Cells were transfected with 5 μg NA plasmids using the ExpiFectamine CHO or ExpiFectamine 293 Transfection Kit (Thermo Fisher Scientific), following the protocol provided with the transfection kit. Transfected cells were then incubated at 37 °C at 8% CO_2 with an orbital shaking speed of 250 rpm (orbital diameter of 25 mm) for 48 h.

Flow cytometry analysis

Selection of cross-reactive plasma against group 1 and group 2 IAV NAs and the evaluation of the breadth of the best FNI mAbs were performed by binding to NA proteins expressed on the cell surface by flow cytometry. NA expressing ExpiCHO-S or Expi293F cells were harvested, washed twice in FACS buffer (PBS supplemented with 2% FBS and 2 mM of ethylenediaminetetraacetic acid), counted and distributed into 96-well U-bottom plates (Corning). The sera or mAbs were serially diluted and incubated with the cells for 45 min on ice and then washed in FACS buffer. Alexa Fluor 647-labelled goat anti-human IgG secondary antibody (Jackson Immunoresearch) was added at 1 μg ml^{-1} to the cells, following 20 min of incubation on ice. Cells were then washed with FACS buffer and analysed

on a ZE5 Flow Cytometer (Bio-Rad). Data were normalized to the NA-positive cell fractions and analysed with FlowJo software.

Pseudoparticles production

NA-only based particles (NA-VLPs) were produced by transfecting HEK293T/17 cells in 56.7-cm^2 cell culture dishes (NUNC). Cells were co-transfected using X-tremeGENE HP (Roche) with 1 µg of NA-expressing plasmid and with 8 µg of a complementing viral-genome reporter vector, pNL4-3.Luc+.E-R+. A 1:3 DNA:X-tremeGENE HP ratio was used and the transfection was performed following the manufacturer's protocol. Transfected cells were incubated for 72 h, then the supernatant was harvested and stored at −80 °C into aliquots. The pseudoparticles used in this study were generated with NAs from H1N1 A/swine Hebei/0116/2017, H1N2 A/swine/Kansas/A02246977/2021, H7N3 A/Ck/Ja/PAVX17170/2017, H7N3 A/Canada/rv504/2004, H6N4 A/mallard duck/Netherlands/30/2011, H6N5 A/aquatic bird/Korea/CN5/2009, H4N6 A/swine Ontario/01911-1/99, H5N6 A/Hangzhou/01/2021, H5N6 A/Ck/Suzhou/j6/2019, H7N7 A/Ck/621572/03, H5N8 A/Ck/Russia/3-29/2020 and H7N9 A/Anhui/1/2013.

Production of recombinant NA proteins

NA proteins containing an N-terminal IgK light chain secretion sequence, a 6×His-Tag, AviTag, a tetrabrachion tetramerization domain and a thrombin or tobacco etch virus (TEV) protease site were expressed in Expi293F cells at 37 °C and 8% CO_2. Cell culture supernatant was collected 4 days post-transfection and proteins were purified using Nickel-NTA agarose (Qiagen). Proteins were buffer exchanged into 50 mM Tris-HCl, 150 mM NaCl and 10 mM $CaCl_2$ (pH 8.0) by dialysis or size-exclusion chromatography using a Superdex 200 10/300 GL Increase or Superdex 200 (16/600) (Cytiva) column. NA from A/Tanzania/205/2010, NA B/Perth/211/2001 and NA B/Malaysia/2506/2004 were buffer exchanged into PBS (136.8 mM NaCl, 2.5 mM KCl, 0.8 mM Na_2HPO_4, 1.47 mM KH_2PO_4, 0.9 mM $CaCl_2$ and 0.5 mM $MgCl_2$, pH 7.4) by dialysis.

ELLA

384-Well (NUNC) or 96-well half-area ELISA (Corning) plates were coated overnight at 4 °C with 25 μg ml^{-1} of fetuin. Plates were blocked with a 1% w/v solution of BSA (Sigma-Aldrich) in PBS containing $Ca^{2+}Mg^{2+}$ (Gibco) and incubated for 1 h at room temperature. In the meantime, mAbs were serially diluted and mixed with a fixed amount of recombinant NA proteins or NA-VLPs and the mix was transferred to fetuin-coated plates, previously washed. After an overnight incubation at 37 °C, the plates were washed and 1 μg ml^{-1} lectin from peanut-agglutinin (Sigma-Aldrich) labelled with horseradish peroxidase was added and incubated at room temperature for 1 h. After further washing, SureBlue Reserve TMB 1-Component (KPL) was added, and plates were read at 450 nm after blocking the reaction with 1% hydrochloric acid solution.

MUNANA

mAbs were serially diluted and mixed with a fixed amount of recombinant NA proteins in 96-well black plates (Greiner) or 384-well black plates (TTP Labtech) and incubated for 30 min at 37 °C. Subsequently, the MUNANA substrate (Sigma Aldrich) was added at a final concentration of 16.7 μM. The plates were incubated for 2.5–5 h at 37 °C, then the reaction was stopped with MUNANA stop solution (0.2 M glycine/50% Et-OH, pH 10.7) and the NA activity was measured by fluorescence using Cytation5 (excitation at 365 nm and emission at 445 nm).

SPR assays to measure binding of Avi-tagged NA proteins by Fabs

Measurements were performed using a Biacore T200 instrument. A CM5 chip with covalently immobilized anti-Avi tag polyclonal antibody (A00674-40, GenScript) was used for surface capture of NA proteins bearing the Avi tag (GLNDIFEAQKIEWHE). Running buffer was filtered with HEPES-buffered saline with added P-20 and calcium chloride (HBS P+ $CaCl_2$: 10 mM HEPES (4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid) pH 7.5, 150 mM NaCl, 0.005% w/v P-20 and 10 mM $CaCl_2$). The regeneration

condition was two injections of 75 mM phosphoric acid for 1 min each. Measurements were performed at 25 °C. Experiments were performed with a fourfold dilution series of monomeric Fabs (800, 200, 50 and 12.5 nM) and were run as single-cycle kinetics.

Data were double-reference subtracted and fit to a binding model using Biacore Insight software. The 1:1 binding model with a constant R_{max} was used to estimate the kinetic parameters. For each ligand (NA protein), a constant R_{max}, which is the maximum of all R_{max} values for all Fab analytes from the default 1:1 binding analysis, was applied, under the assumption that all Fabs have the same R_{max} for a specific NA protein ligand. The consistent capture level of each ligand over all cycles (within 8%) justifies this assumption. The experiment was performed twice with two technical replicates for each ligand (NA protein). K_d values were reported as the average of two replicates. For measurements in which the dissociation constant K_d is slower than the detection range of the instrument ($K_d < 10^{-6}$ s^{-1}), the binding affinity (K_d) is denoted as less than 0.01 nM.

Virus titration and in vitro neutralization assays

Virus titres were determined by FFU assay on MDCK-LN cells, using immunofluorescence staining. MDCK-LN cells were seeded into flat-bottom tissue culture 96-well plates (3904, Corning) at 25,000 cells per well and cultured overnight at 37 °C in growth medium. Twenty-four hours later, fivefold dilutions of viral stocks were prepared in infection medium containing 0.5 μg ml^{-1} TPCK-treated trypsin, with each dilution tested in quadruplicate. Cells were washed with DMEM medium and then infected with 50 μl of diluted virus. After absorption at 37 °C for 30 min, inoculum was removed, cells were washed once in DMEM medium and 100 μl infection medium with 2.4% colloidal cellulose overlay (435244, Sigma) and 0.5 μg ml^{-1} TPCK-treated trypsin was added to each well. After 24 h of incubation at 37 °C for IAV and 35 °C for IBV, cells were fixed with 4% paraformaldehyde (AA233689L, Fisher) for 30 min at room temperature, washed three times with wash/perm buffer (0.5% Triton X-100 in PBS) to remove residual paraformaldehyde, then the cells were permeabilized with 100 μl wash/perm buffer for 30 min. Cells were incubated for 1 h with 50 μl

anti-influenza A or anti-influenza B nucleoprotein antibody (ab128193 or ab20711, Abcam) at 1:1,000 and 0.5% BSA in wash/perm buffer followed by three washes with wash/perm buffer. The cells were then incubated with 50 µl goat anti-mouse IgG Alexa Fluor Plus 647 (A-32728, Invitrogen) at 1:1,000 with 2 µg ml^{-1} Hoechst dye (10 mg ml^{-1}; H3570, Invitrogen) in wash/perm buffer for 1 h and then three additional washes with wash/perm buffer. Plates were imaged on a S6 Universal M2 Plate Imager and resulting FFU titres were calculated by counting wells that contained 10–20 distinct foci.

Microneutralization was evaluated in a multicycle assay by adapting a protocol previously described[55]. In brief, MDCK-LN cells were seeded into flat-bottom tissue culture 96-well plates at 25,000 cells per well and cultured overnight at 37 °C in growth medium. Twenty-four hours later, virus stocks were diluted in infection medium containing 0.5 µg ml^{-1} TPCK-treated trypsin for a final concentration of 70 FFU per well, cells were washed once with DMEM medium, then 50 µl of diluted virus was added to the cells and incubated at 37 °C for 30 min. A nine-point 1:5 serial dilution of mAb was prepared in infection medium containing 0.5 µg ml^{-1} TPCK-treated trypsin, with each dilution tested in triplicate (final range: 50,000–0.128 ng ml^{-1}). Diluted antibody was mixed 1:1 with colloidal cellulose overlay (final concentration of 2.4%) resuspended in infection medium containing 0.5 µg ml^{-1} TPCK-treated trypsin, then infected cells were washed once with DMEM medium and 100 µl antibody with cellulose overlay was added to the wells. Cells were incubated for 24 h at 37 °C for IAV, 35 °C for IBV B/Victoria/2/87-like viruses and 35 °C for 48 h for IBV B/Yamagata/16/88-like viruses in a tissue culture incubator. Viral replication was measured following the same procedure as described above for the virus titration assay. Plates were imaged on a Cytation5 plate reader. Whole-well images were acquired (12 images at ×4 magnification per well) and nucleocapsid-positive cells were counted using the manufacturer's software. To control for background, the mean count of positive cells in uninfected wells was calculated and subtracted from all data points. All positive cell counts were normalized to the no-antibody treatment control. Data were analysed and graphed using GraphPad Prism software (v9.0.0). The IC_{50} values were calculated using a non-linear regression model (variable slope model, four parameters) of log(inhibitor) versus response and the IC_{50} values were

interpolated from the curve at $y = 50$. The geometric mean of two independent experiments in which the antibody was tested in triplicate was calculated in Excel.

To confirm the actual $TCID_{50}$ of the input virus used in the neutralization assay for each viral strain, a $TCID_{50}$ viral assay was carried out in parallel. For this assay, a twofold serial dilution of input virus in 50 µl was added to duplicate wells containing MDCK-LN cells and incubated in parallel to the microneutralization assay at 37 °C for 30 min. Cells were washed once with DMEM, and 100 µl infection medium containing 0.5 $\mu g\,ml^{-1}$ TPCK-treated trypsin was added to the cells and incubated with the neutralization plates. Viral replication was measured following the same procedure as described above for the virus titration assay. Infected wells were observed with the Cytation5 and resulting titres were calculated using the Reed–Muench method[56]. The target viral input for each experiment was 100 $TCID_{50}$ per well.

Assessment of synergy

Cell plating and virus infections were carried out using the in vitro virus neutralization procedure described above with all infection medium containing 0.5 $\mu g\,ml^{-1}$ TPCK-treated trypsin. To assess the synergy between the two monoclonal antibodies, a seven-point 1:3 serial dilution of MEDI8852 was prepared in infection medium (final concentration: 10–0.01372 $\mu g\,ml^{-1}$). Subsequently, a nine-point 1:3 serial dilution of FNI9 was prepared in infection medium (final concentration: 7.5–0.001143 $\mu g\,ml^{-1}$) and mixed at 1:1 ratio with MEDI8852 in 2-ml deep-well plates creating a drug combo checkerboard matrix. The diluted drug combination was then mixed 1:1 with a colloidal cellulose overlay (final concentration of 2.4%). The infected cells were washed once with DMEM and 100 µl of antibody–drug combination with cellulose overlay was added to the wells, with each condition tested in triplicate. The plates were incubated at 37 °C for 24 h in a tissue culture incubator. Viral replication was measured following the same procedure as described above in the virus titration assay and plates were imaged on the Cytation5 plate reader. Neutralization synergy was calculated by normalizing the signal of the treatment wells to no-antibody virus control

wells and no-virus cell control wells and the values were imported to the web-based software SynergyFinder. The zero interaction potency statistical model was utilized to calculate total synergy scores, with values greater than 10 indicating likely synergy between the two drugs. Two independent experiments were performed with three technical replicates for each virus.

In vitro resistance studies

Antibody-resistant mutants were isolated by screening the viral population in bulk or by serial passaging the virus in the presence of an escalating dose of antibody. For the former approach, MDCK cells were plated into 96-well plates (CON3596, Corning) at 30,000 cells per well in 100 µl of culture medium (MEM + GlutaMAX (41090-028, Thermo Fisher) + 10% FBS Hyclone (SH30070.03, VWR) + 50 U ml^{-1} penicillin–50 ng ml^{-1} streptomycin (4-01F00-H, Bioconcept) + 100 µg ml^{-1} kanamycin (15160047, Life Technologies) + 1% NEAA (5-13K00-H, Bioconcept) + 1% sodium pyruvate (11360039, Life Technologies) + 0.05 mM β-mercaptoethanol (5-69F00-E, Bioconcept)) and cultured overnight at 37 °C at 5% CO_2. The day after, A/Hong Kong/1/1968 virus input was pre-incubated with 100 µg ml^{-1} of FNI19 antibody for 30 min at 37 °C. The virus–antibody mix was then limiting diluted in infection medium (MEM supplemented with 1 µg ml^{-1} TCPK-trypsin (LS003750, Bioconcept) and 10 µg ml^{-1} kanamycin) to infect MDCK cells at MOI of 0.001 in 100 µl, after two washing steps with 200 µl per well of PBS (D8537-500ML, Sigma-Aldrich). Three hours later, 100 µl per well of infection medium containing 100 µg ml^{-1} FNI19 was added, and cells were incubated for a further 72 h. After the incubation time, a 20 mM MUNANA substrate (69587-5MG, Sigma) solution was prepared in MUNANA buffer (32.5 mM MES and 4 mM $CaCl_2$, pH 6.5) and 50 µl per well dispensed into 96-well plate to be incubated with an equal amount of assay culture supernatant for 1 h at 37 °C. Of stop solution (0.2 M glycine and 50% Et-OH, pH 10.7), 100 µl per well was added and fluorescence at 445 nm was measured with a Cytation5 (Biotek-Agilent). Fluorescence threshold was set to 5× the signal of not infected cells and supernatants above that cut-off were processed for NA sequencing. Viral genomic RNA was extracted from 140 µl of culture supernatant using the Qiagen QIAmp Viral RNA kit (52904, Qiagen)

according to manufacturer's instruction. cDNA was synthetized using the SuperScript III RT kit (18080044, Life Technologies) with the primer Uni12 (5′-AGC RAA AGC AGG-3′). The gene encoding NA was amplified by PCR using the Q5 HotStart DNA polymerase kit (M0493L, Bioconcept) with a specific primer forward: 5′-caattggctctgtctctct-3′ and reverse: 5′-atgaaattgatgttcgccc-3′. PCR product was purified from 1.5% agarose gel and sequenced with the three different primers: 5′-caattggctctgtctctct-3′, 5′-gaagagccgatactagaa-3′ and 5′-ttctagtatcggctcttc-3′. Sequence alignment with A/Hong Kong/1/1968 NA was performed with CLC Main Workbench 22 software.

For the serial passaging resistance experiment, MDCK-LN cells were seeded in 24-well plates at 120,000 cells per well and cultured overnight at 37 °C in growth medium (DMEM, 10% FBS, 0.01 M HEPES and 100 U ml^{-1} penicillin–100 μg ml^{-1} streptomycin). Twenty-four hours later, H1N1 A/California/07/2009 or A/New Caledonia/20/99 virus stock and FNI9 were diluted and mixed 1:1 in infection medium containing 0.5 μg ml^{-1} TPCK-treated trypsin for a final concentration of 1,000 PFU per well virus and 0.5× IC_{50} of FNI9 (47.35 ng ml^{-1}). The virus–antibody mixture was placed at 37 °C and precomplexed for 30 min. Cell and virus control wells were included for cytophatic effects (CPE) comparators. The cells were washed twice in DMEM and then 100 μl of the virus–antibody complexed mixture, virus and cell controls were added to the wells for 1 h at 37 °C. After virus adsorption, the cells were washed twice in DMEM, then 500 μl of diluted antibody or infection medium containing 0.5 μg ml^{-1} TPCK-treated trypsin was added back to their respective wells. The plates were incubated at 37 °C and the virus was collected once there was over 50% CPE in the virus–antibody well. The virus supernatant was frozen in 120 μl aliquots at −80 °C, one aliquot was used for the next resistance passage, one aliquot was added to 360 μl of DNA/RNA Shield for RNA extraction, and one aliquot was saved for additional propagations. For passage 2, the process was repeated with the original 0.5× and a 1× IC_{50} concentration of FNI9, with the highest concentration of antibody showing CPE being selected for the next passage. For each additional passage, the concentration from the selected well was repeated along with a 2× concentration. For the CA09 study, there were eight passages of virus and a final concentration of 64× the IC_{50} of FNI9 (3,200 ng ml^{-1}).

For identification of genomic mutations, viral genomic RNA was extracted using the Zymo QuickRNA Viral Kit (R1034) according to the manufacturer’s instructions and the RNA concentration was measured using the Qubit RNA BR Assay Kit (Q10210, Invitrogen). cDNA was synthesized using the SuperScript III One-Step RT–PCR system with Platinum Taq DNA polymerase (12574018, Invitrogen) using the MBTuni-12 (5′-ACGCGTGATCAGCAAAAGCAGG-3′) and MBTuni-13 (5′-ACGCGTGATCAGTAGAAACAAGG-3′) primers, and the PCR fragments were purified using AMPure XP magnetic beads (A63880, Beckman Coulter). Next, the amplicons were fragmented using the NEBNext Ultra II FS DNA Library Prep Kit for Illumina (E6177L, NEB), then adaptors and indexes were added using NEBNext Multiplex Oligos for Illumina (7335L, NEB). DNA samples were purified using NEBNext Sample purification beads and then quantified using the Qubit DNA HS Assay kit (Q32854, Invitrogen), and verified using a D500 tape station (G2964AA, Agilent). Finally, the samples were diluted and pooled, then sequenced using a MiSEQ analyser with the MiSeq Reagent Kit v3 (MS-102-3003) and PhiX Control v3 (FC-110-3001). The NA mutations I223R and S247N first appeared in passages 6 and 7, respectively, and each reached 50% of the total next-generation sequencing (NGS) reads by passage 8.

Measurement of Fc-effector functions

To evaluate complement-dependent cytotoxicity (CDC assay), MDCK-LN cells were infected with A/California/07/2009 (H1N1) at an MOI of 6. After 18 h at 37 °C with 5% CO_2, cells were detached with trypsin, washed and dispensed into flat-bottom 384-well plates. Antibodies were serially diluted in AIM-V medium, mixed with infected target cells and incubated for 10 min at room temperature. After the incubation, guinea pig low tox complement (Cedarlane) previously diluted 1:5 with AIM-V medium was added to each well. Control wells containing target cells and complement with 50 µl of 2% Triton X-100 to measure maximal lysis or target cells and complement only to evaluate spontaneous lysis were also included. Cell death was quantified after 3 h of incubation at 37 °C with 5% CO_2, by measuring lactate dehydrogenase (LDH) release using the LDH detection kit (11644793001, Roche) according to the manufacturer’s instructions. Using a

kinetic protocol, the absorbance at 490 nm and 650 nm was measured once a minute for 8 min, and the percent specific lysis was determined.

Antibody-dependent cell cytotoxicity was performed on A549 cells infected with A/Puerto Rico/8/34 (H1N1) at an MOI of 6. After 18 h at 37 °C with 5% CO_2, infected target cells were detached with trypsin, washed and dispended into round-bottom 384-well plates (7.5×10^3 per well in AIM-V). The serial dilutions for antibodies were prepared in AIM-V, added onto cells, and incubated for 10 min at room temperature. Human natural killer (NK) cells were isolated from fresh blood using the MACSxpress WB NK cell isolation kit, human (130-127-695, Miltenyi Biotec) and following the manufacturer's instructions. After the incubation, NK cells were added at a cell density of 4.5×10^4 per well, to reach an effector to target ratio of 6:1. Control wells to measure maximal lysis (containing target cells with 3% Triton X-100) and spontaneous lysis (containing target cells and effector cells without antibody) were also included. Cell death was determined after 4 h of incubation at 37 °C with 5% CO_2 by measuring LDH release using a LDH detection kit (11644793001, Roche) according to the manufacturer's instructions. Using a kinetic protocol, the absorbance at 490 nm and 650 nm was measured once every 2 min for 8 min, and the percent specific lysis was determined. Determination of antibody-dependent cellular phagocytosis was performed using ExpiCHO transiently expressing NA from A/Perth/16/2009 (H3N2) and labelled with PKH67 (Sigma-Aldrich). Peripheral blood mononuclear cells from a healthy donor were labelled with CellTrace Violet (Invitrogen) and used as a source of phagocytic effector cells. Antibodies were serially diluted in RPMI-1640 medium. Labelled target cells (1×10^4 per well) were incubated with the diluted mAbs for 10 min and then mixed with labelled PBMCs (1.8×10^5 per well). Following an overnight incubation at 37 °C and 5% CO_2, cells were stained for 20 min with APC-labelled anti-CD14 mAb (BD Pharmingen), BV605-labelled anti-CD16 mAb (BioLegend), BV711-labelled anti-CD19 mAb (BioLegend), PerCP/Cy5.5-labelled anti-CD3 mAb (BioLegend) and APC/Cy7-labelled anti-CD56 mAb (BioLegend) for the identification of $CD14^+$ monocytes. After the incubation, cells were washed and fixed with 4% paraformaldehyde before acquisition on a ZE5 Cell Analyzer (Bio-Rad). Data were analysed using FlowJo software. The percentage of antibody-dependent cellular phagocytosis was calculated as the percentage of

monocytes ($CD14^+$ cells) positive for PKH67. The assays were performed on two different donors and the area under the curves were calculated.

Cryo-EM sample preparation

A summary of sample preparation is provided in Supplementary Table 5. In brief, NA proteins were mixed with a 1.3–1.5-fold molar excess of FNI Fabs and incubated on ice for 1 h. Complexes were purified by size-exclusion chromatography using a Superdex 200 10/300 GL Increase (Cytiva) column equilibrated in 50 mM Tris-HCl (pH 8.0), 150 mM NaCl and 10 mM $CaCl_2$. Complex containing fractions were pooled and concentrated using Amicon Ultra centrifugal filters (Millipore Sigma) before vitrification. Of 0.2–0.25 mg ml^{-1} of each NA–Fab complex with or without *n*-dodecyl-β-maltoside (DDM) at 0.70× critical micelle concentration (CMC), 3 µl were loaded onto a freshly glow discharged 1.2/1.3 UltrAuFoil grids before plunge freezing using a Vitrobot MarkIV (Thermo Fisher Scientific) with a blot force of 3 and a blot time of 7 s at 100% humidity and 4 °C.

Cryo-EM data collection

The FNI9–NA (N2 A/Tanzania/205/2010) dataset was collected on a 300 kV Titan Krios equipped with a Gatan K3 direct detector and Gatan BioQuantum energy filter, operated with a slit width of 30 eV. Semi-automated data collection was carried out using SerialEM[57] at a nominal magnification of ×81,000 in super-resolution mode giving a pixel size of 0.53 Å. The dose was 48 e^- $Å^{-2}$, fractionated over 40 frames of 75 ms each.

For other datasets, data were acquired on either a 200 kV FEI Glacios transmission electron microscope (TEM) equipped with a Falcon 4 camera, or a 300 kV FEI Titan Krios TEM equipped with a Gatan K3 Summit direct detector and Gatan Quantum GIF energy filter, operated in zero-loss mode with a slit width of 20 eV. Automated data collection was carried out using Leginon[58] at a nominal magnification of ×150,000 and a pixel size of 0.927 Å for the Glacios datasets, and ×105,000 magnification and a pixel size of 0.823 Å for the Krios datasets. The dose rate was adjusted to 1.4 counts per pixel per second, and each movie was fractionated in 24 frames for 110 ms for Glacios datasets, and 35 or 38 frames for 40 ms for the Krios datasets.

For each dataset, between 1,089 and 5,433 movies were collected with a defocus range comprised between −1.0 and −2.5 μm. A summary of data collection is provided in Supplementary Table 5.

Cryo-EM data processing

For the FNI9–NA (N2 A/Tanzania/205/2010) dataset, Relion[59,60] was used for cryo-electron microscopy (cryo-EM) data processing. Dose-weighted movie frame alignment was done using a Relion implementation of MotionCor2 (ref. [61]) to account for stage drift and beam-induced motion. The contrast transfer function (CTF) was estimated for each micrograph using CTFFIND4 (ref. [62]). A Laplacian-of-Gaussian algorithm was used for template-free automated particle picking. Extracted particles were subjected to 2D classification and only those contributing to detailed views of a NA–Fab average were used for further analysis. An initial 3D map was generated ab initio from all selected particles using Relion. An NA–Fab complex with four Fabs bound was used as input to 3D classification where 3-Fab and 4-Fab complexes were separated. These separate 3D classes were submitted to further refinement that included CTF refinement, Bayesian polishing and masking. No symmetry (C1) was applied for the further refinement of the NA–3-Fab map, but C4 symmetry was applied for the NA–4-Fab refinement.

For all other datasets, dose-weighted movie frame alignment was done using MotionCor2 (ref. [61]) or full-frame or Patch motion correction in cryoSPARC[63] to account for stage drift and beam-induced motion. The CTF was estimated for each micrograph using CTFfind4 (ref. [62]), gCTF[64] or Patch CTF in cryoSPARC. Individual particles are selected using automated picking protocols and extracted into particle stacks in either Relion[59,60] or cryoSPARC. Extracted particles and/or 2D classes showing intact particles from Relion 3.0 were subjected to 2D and/or 3D classification in cryoSPARC. Initial maps were generated ab initio from all selected particles. The best 3D classes were submitted to homogeneous 3D refinement that included dynamic masking and either C1 (for 1-Fab or 3-Fab bound NAs) or C4 (for 4-Fab bound NAs) symmetry applied.

Reported resolutions are based on the gold-standard Fourier shell correlation of 0.143 criterion[65]. A summary of cryo-EM data statistics is provided in Supplementary Table 5.

Cryo-EM model building and analysis

Atomic model of N2 NA from A/Tanzania/205/2010 H3N2 (Protein Data Bank (PDB) ID: 4GZX) or NA from B/Phuket/3073/2013 (PDB: 6V4O) were used as the initial model to fit into cryo-EM maps by UCSF Chimera[66]. Coot[67] and ISOLDE[68] were used to manually rebuild amino acid mutations and glycans on NAs and build the Fabs. Models were self-restrained by ProSMART[69] and refined by Refmac Servalcat[70] in the ccpem v1.6.0 software suite[71]. Model validation and map-model FSC were generated by validation: model tool in ccpem v1.6.0.

Figures were generated using Pymol[72]. Epitopes in all NA complexes were identified by determining NA residues within 5.0 Å of any atoms in FNI Fabs, 1G01, SA or oseltamivir using MOE (v2020.0901; https://www.chemcomp.com) contact analysis with default settings. MOE QuickPrep was used to prepare the NA complexes for static epitope analysis.

Glycan profiling of NA with peptide mapping LC–MS

Peptide mapping with liquid chromatography–mass spectrometry (LC–MS) was used to profile the site-specific glycosylation sites on two NAs (A/Tanzania/205/2010 and A/Hong Kong/2671/2019). Glycopeptides containing only one specific glycan were achieved by selectively digesting with trypsin, Glu-C, Lys-C or Asp-N protease, depending on the sequence context. Of each digest product (peptide with a single glycan), 25 µg was analysed by LC–MS (Agilent AdvanceBio peptide mapping column and Thermo Q Exactive Plus Orbitrap MS). Peptide mapping data were analysed on Biopharma Finder 3.2 data analysis software. Technical replicates were performed by injecting 25 µg of digested product three times from the same sample vial into the LC–MS.

MD methods

The coordinates of FNI17–NA (Tanzania/2010) and FNI9–NA (Tanzania/2010) were obtained from cryo-EM (see above) with glycans determined by peptide mapping LC–MS (Extended Data Fig. 7) modelled using ISOLDE as previously described[68,73]. The resulting model was then prepared using QuickPrep (MOE v2020.0901; https://www.chemcomp.com).

The FNI9–NA tetramer and FNI17–NA tetramer complex structures were each parameterized for MD using tleap in AMBER[74] with the ff14SB protein force field[75], the GLYCAM_06j-1 glycan force field[76] and the TIP3P water force field[77]; the Joung and Cheatham force field[78] was used for ions.

We generated 11.5 µs of aggregate MD for each of the two FNI–NA complexes—23 µs in total—by performing ten independent MD simulations each seeded with different initial velocities. Each of these 20 simulations was minimized and equilibrated as previously described[79]. Specifically, all heavy atoms resolved in the experimental data were restrained: during minimization, 100 ps of MD heating to 300 K; 100 ps of MD at 300 K; 250 ps of MD with a tenfold weaker restraint force constant at 300 K. With all backbone atoms resolved in the structure restrained: 10,000 steps of minimization; 100 ps of MD at 300 K; 100 ps of MD at 300 K (tenfold weaker restraint force constant); 100 ps of MD at 300 K with a further tenfold reduced force constant; 100 ps of MD at 300 K with another tenfold reduced force constant (0.1 kcal per mol Å^2). Finally, 2.5 ns of unrestrained MD at 300 K. Production MD involved 0.8 µs of unrestrained MD simulation for each of the 20 independent simulations, followed by an additional 0.7 µs for five of the FNI9–NA and five of the FNI17–NA simulations for an aggregate of 11.5 µs of MD for FNI17–NA and 11.5 µs of MD for FNI9–NA.

Simulations were processed using cpptraj[80] as previously described[79]. Epitope–paratope interactions were computed using contact analysis in MOE (CCG MOE 2020.09) in which all hydrogen bond, metal, ionic, arene and distance-based interactions (within 5 Å) were accounted for. This analysis was performed on MD frames sampled every 10 ns. The MOE energy scores reported are the average energy for each interaction pair in the aggregate MD. The percent occupancy reported reflects the percentage of

frames in the 11.5-µs MD simulation datasets in which a particular contact was present.

In vivo studies

Prophylactic efficacy of muFNI9 and muFM08 in BALB/C mice infected with A/Puerto Rico/8/34 was assessed at Burleson Research Technology Inc. (North Carolina) following approval of animal procedures by the Testing Facility's Institutional Animal Care and Use Committee (IACUC). Seven-to-eight-week-old BALB/c female mice (Charles River Laboratories) received a single dose of muFNI9, muMEDI8852, muFNI9(N297Q), muMEDI8852(N297Q) or vehicle control on day −1 via intravenous injection, 24 h before infection, based on day −2 individual body weight. On day 0, animals ($n = 6$ per group) were infected intranasally with influenza A/Puerto Rico/8/34 virus (VR-95PQ, ATCC). Mice were monitored once daily from days 0–3, and twice daily from days 4–14. Body weight was measured daily from day 4 to day 14, and serum was collected at day 0 (pre-infection) for mAb quantitation.

The prophylactic activity of MEDI8852 and FNI17 mAbs administered singly or in a 1:1 combination on the response to H1N1 A/Puerto Rico/8/34 infection or human MEDI8852, FNI19 and muMEDI8852 and muFNI9 mAbs on the response to H3N2 A/Hong Kong/1/1968 was assessed in BALB/c female mice at 0.25 or 0.125 mg kg^{-1}. Body weight loss and survival from days 0 through to day 14 post-infection were used as major end point. The area of the negative peaks, defined as the area between the body weight line of each animal over the 14 days and the 0% body weight loss baseline, was calculated using GraphPad prism 9 and used to compare the prophylactic activity of the anti-HA stem, anti-NA and the combination of the two mAbs.

The prophylactic efficacy of FNI9 to influenza B/Victoria/504/2000 and B/Brisbane/60/2008 was assessed in **7**–8-week-old BALB/c female mice by administering a single dose of FNI9 or vehicle control on day −1 via intravenous injection, 24 h before infection. Mice in the oseltamivir (MedChemExpress) treatment groups received 10 mg kg^{-1} of small drug via oral gavage 2 h pre-infection, 6 h post-infection and then once daily until day

3 post-infection based on day −2 individual body weight. Animals were infected intranasally with influenza B/Victoria/504/2000 or influenza B/Brisbane/60/2008 (ViraPur) virus on day 0. Serum was collected 2 h before infection on day 0 for IgG quantitation. Animals were euthanized on day 3 and lung and body weights were recorded. Infectious virus titres were measured in lung homogenate supernatants collected on day 3 (PFU per gram of lung) using a plaque assay.

The in vivo prophylactic studies involving human FNI9 IgG mAb and A/Puerto Rico/8/34 (PR8) and A/Singapore/INFHH-16-19/2016 viruses were performed in compliance with federal laws and institutional guidelines and have been approved by the Washington University in St. Louis IACUC. Seven-to-nine-week-old female BALB/c mice were randomized based on age before antibody treatment and virus infection. Anti-NA mAbs were administered as a single intravenous injection 24 h before virus infection. For oseltamivir treatment, the mice were orally dosed with 10 mg kg^{-1} of oseltamivir phosphate 2 h before infection and then once daily for 5 days. On the day of infection, mice were anaesthetized with 2% isoflurane and intranasally infected with 5 times the lethal dose 50 (5 LD_{50}) of viruses diluted in 50 µl of PBS. After infection, mice were monitored daily, and their weights were recorded. To determine lung viral titres, mice were killed 4 days post-infection. Lungs were harvested and stored at −80 °C. Tissues were homogenized in 1 ml of MEM, and serial dilutions of homogenates were incubated on MDCK-LN cells for 1 h at 37 °C followed by two PBS washes. Of infection medium (MEM with 100 U ml^{-1} penicillin, 50 µg ml^{-1} streptomycin, 100 µg ml^{-1} kanamycin and 0.5 µg ml^{-1} TPCK-treated trypsin), 200 µl was added to the cells. After 3 days of incubation at 37 °C, 50 µl of supernatants were removed and added to 50 µl of 0.5% turkey red blood cells and the end point dilution of haemagglutination was scored. $TCID_{50}$ titres were calculated using the Spearman–Karber formula.

IgG quantification was performed using an electrochemiluminescence-based immunoassay. Standard curve and quality controls were diluted in BALB/C mouse serum (Biovit) and stored at −20 °C in aliquots until needed for the quantitation assay. Reagents and samples were equilibrated to room temperature for at least 30 min. Sera, standard curve and QCs were then centrifuged at 1,000*g* for 5 min before use. For the murine mAbs

quantification, multi-array 96-well standard plates (Meso Scale Discovery) were coated with NA protein (N9) or HA protein (H5) at 4 μg ml^{-1} or 2 μg ml^{-1}, respectively. Plates were incubated overnight at 4 °C, then washed with PBS–0.05% Tween (PBS-T) and blocked with casein in PBS (Thermo Fisher) at room temperature for 2 h. Sera were diluted 1:20 with BALB/C mouse serum, followed by a 1:50 (for anti-NA mAb) or 1:100 (for anti-HA mAb) dilution with casein, that was performed for standard curve and QCs as well. Plates were washed with PBS-T, then 50 μl of diluted samples, standard curve and QCs were added to the appropriate wells and incubated for 1 h at room temperature. Goat anti-mouse-SULFO TAG (Meso Scale Diagnostics) was used for detection at a concentration of 1 μg ml^{-1} for FNI9 and 0.75 μg ml^{-1} for muMEDI8852 to the plates previously washed with PBS-T and incubated for 1 h. After final washes with PBS-T, 150 μl of MSD Read buffer was added per well, and plates were read on an MESO Quickplex SQ 120. Human mAbs were quantified with the same method but using an anti-LS antibody for capture and anti-CH2-SULFO TAG for detection.

Deep mutational scanning profiling

A deep mutational scanning approach was used to determine how NA amino acid mutations affect mAb binding. A cell-surface display library based on N1 A/California/07/2009 was generated with all variants for 170 select amino acids included. The mutated residues include epitope residues within 8.0 Å distance from bound Fabs obtained from cryo-EM (FNI17–Fab and FNI19–Fab in complex with N2 A/Tanzania/205/2010; this study) and crystal structures (FNI3–Fab with N2 A/Tanzania/205/2010 (unpublished), 1E01–Fab (PDB: 6Q20), Mem5–Fab (PDB: 2AEP) and B10–Fab (PDB: 6N6B)). In addition, residues with known sequence variation in GISAID in the past 20 years for H1N1 and H5N1 were included.

Library construction, lentivirus production, next-generation sequencing and data deconvolution were performed by the Genetic Perturbation Platform (GPP) of the Broad Institute of MIT and Harvard. A detailed description of each of these methods has been previously described[81]. For library synthesis, the lentiviral vector pMT025 encoding the wild-type N1 A/California/07/2009 open reading frame with the addition of a C-terminal

minimal HA tag via a 2×G_4S linker was synthesized and used as a template to generate the variant library pool. The full-length variant pool with 5′ and 3′ flanking adapters containing NheI and MluI, respectively, was synthesized at Twist Biosciences. Following restriction digest cloning and pDNA amplification, the variant composition of the pDNA library was assessed by NGS using the Illumina Nextera XT platform. All 3,472 designed variants were detected, and the distribution of the variants was approximately log-normal with a 1.3-fold standard deviation in the read counts across the variants. Lentiviral particles were produced according to protocols available through the GPP (http://www.broadinstitute.org/rnai/public/resources/protocols). The library was transduced into FreeStyle 293-F Cells (Thermo Fisher Scientific) at a MOI of approximately 0.3 and maintained at an average representation of 2,000 cells per variant.

For the screen, mAb binding to cell-surface-displayed NA was determined by flow cytometry. mAbs were directly conjugated to Alexa Fluor 647 (Thermo Fisher Scientific) through free amine conjugation to enable fluorescence-based detection. 293-F cells containing the NA variant library were harvested 5–7 days after puromycin selection, washed in PBS containing 3% BSA and then incubated with Alexa Fluor 647-conjugated mAbs for 1 h at 4 °C in the dark. Cells were washed and fixed with 4% paraformaldehyde, then sorted in technical duplicate on two separate sorters (BD Aria Fusion or Sony MA900). Dead cells were excluded from the analysis using the LIVE/DEAD Fixable Blue Dead Cell Stain Kit (Thermo Fisher Scientific). A minimum of 60 million cells were prepared for FACS per technical replicate. 293-F cells expressing the NA variant library were sorted by FACS into negative, low and high mAb-binding bins, and the distribution of variants in each bin was determined by NGS as described below. The NA open reading frame includes a minimal C-terminal HA tag, which was also screened against using the Alexa Fluor 488 anti-HA.11 epitope tag antibody (BioLegend) to determine how mutations at each of the 170 selected residues affect general protein expression. For each screen, a minimum of 5 million unsorted library cells were collected to serve as unsorted controls.

Genomic DNA was isolated from sorted samples and unsorted controls with the QIAamp DNA FFPE Tissue Kit (with the following modification: 0.3 M

NaCl was added to the lysis buffer and incubation at 56 °C extended to 4 h) and the QIAamp DNA Blood Kit (Qiagen), respectively. Genomic DNA was prepared for NGS as previously described[81]. In brief, the open reading frame was amplified from genomic DNA by PCR with the following primers (forward: 5′-ATTCTCCTTGGAATTTGCCCTT-3′ and reverse: 5′-CATAGCGTAAAAGGAGCAACA-3′) and processed for NGS using the Illumina Nextera XT kit. Samples were sequenced using the Nextseq2000 sequencing platform to obtain paired-end reads with 2 × 150 cycles. Sequence data processing and variant calling using the mutant detection software ASMv1.0 was performed as previously described[81]. Normalized reads for each sample are $\log_2$ transformed. $\log_2$ fold-change values comparing the negative sorted bin to the unsorted pool and the low-binding and high-binding FACS-sorted bins for each library variant were computed and then averaged both for the anti-HA tag and the anti-NA antibody. Potential escapes were defined as those having an HA tag signal close to the silent mutations (maximum of two standard deviations away from the mean of the silent mutations) and having an anti-NA signal close to the stop codon mutations (maximum of three standard deviations away from the mean of the stop codon mutations).

Sequence conservation analysis

H3N2 (IAV), H1N1 (IAV), H5N1 (IAV), H7N9 (IAV), H5N8 (IAV), H5N6 (IAV), Victoria (IBV) and Yamagata (IBV) NA protein sequences were retrieved from GISAID (www.gisaid.org). Protein sequences were aligned to a reference NA sequence using MAFFT[82]. For Fig. 2, the sequences used were retrieved from the GISAID repository from January 2000 to October 2022, and the respective reference used for alignment was A/California/07/2009 (NC_026434.1) for H1N1 sequences, A/NewYork/392/2004 (YP_308842.1) for H3N2 sequences and B/Yamagata/16/1988 (AAN39803.1) for Victoria and Yamagata IBV sequences. For Extended Data Fig. 5, the sequences used were retrieved in July 2022 and the respective reference used for alignment was A/NewYork/392/2004 (YP_308842.1) for H3N2, H1N1, H5N1, H7N9 H5N8 and H5N6 sequences, and B/Yamagata/16/1988 (AAN39803.1) for Victoria and Yamagata IBV sequences. The multiple sequence alignments were analysed with R (https://www.R-project.org/) v.4.0.4. The logo plots

were generated with the R package ‘ggseqlogo’ v.0.1 (ref. [83]). The conservation per residue was computed with the R package ‘Biostrings’ v.2.58.0 (https://bioconductor.org/packages/Biostrings).

Statistical analysis

All statistical tests were performed as described in the indicated figure legends using Prism v9.0. The number of independent experiments performed is indicated in the relevant figure legends.

Material availability

Materials generated in this study can be available on request from the corresponding authors and may require a material transfer agreement.

Ethical statement

All experiments involving animals reported in Fig. 5a–d,g,h and Extended Data Fig. 9 were approved by the Burleson Research Technologies, Inc. (BRT) Animal Care and Use Committee (study number: BRT20220306, BRT 2022072 and BRTQ20220307) and by the Office of Laboratory Animal Welfare (PHS assurance number D16-00898). The BRT has been accredited by the Association for Assessment and Accreditation of Laboratory Animal Care (AAALAC) International. The IACUC chair who approved the protocols was also the Lead Veterinarian for these studies, and together with the Study Director was kept apprised of body weight loss and clinical observations for these studies daily. In addition, both the Lead Veterinarian and the Study Director personally checked on the status of study mice throughout each of the studies.

On the basis of extended experience with the influenza model at the BRT, specific weight loss limits were not set for the study reported in Fig. 5a–d and Extended Data Fig. 9 (BRT20220306 and BRTQ20220711), as mice with 30% (or greater) body weight loss had normal cage activity levels indicated by the ability to climb to access the food hoppers and build nests and to interact with cage mates as expected. In addition, mice with weight loss that exceeded 30% were shown to be able to recover from infection in

this model. In such scenarios, the BRT would not euthanize the mouse but instead would wait until the next observation period to reassess behaviour along with any further body weight loss. At each observation period, both the Lead Veterinarian and the Study Director specifically checked for animals that met euthanasia requirements as listed in both the study protocol and the IACUC protocol.

Given the nature of the study and the weight loss involved, measures were undertaken to minimize pain, distress and discomfort as much as possible, including providing DietGel to aid in providing hydration and nutrition from the time of infection through to day 14, increased frequency of animal observations (twice daily, at least 5 h between observations on days 4–14 and on other days if adverse clinical signs were observed), daily body weight and clinical observation updates provided to the Study Director and veterinarian, and humane euthanasia if animals were deemed unlikely to recover from infection.

In addition to protocol-driven euthanasia requirements, although body weight loss was not used as an end point, it was used to further evaluate the mice and help to inform euthanasia decisions. If, for example, body weight loss was high and was not stabilizing and/or if body weight loss was high and other clinical signs (especially a strong decrease in movement or activity) were observed, animals were immediately humanely euthanized as either scenario was generally a good indication that the mouse was unlikely to recover from infection.

Procedures, facilities and housing were in compliance with the NRC Guide for the Care and Use of Laboratory Animals. Mice were euthanized at 3 days post-infection (Fig. 5g,h; BRT 2022072) or throughout a 14-day observation period (Fig. 5a–d and Extended Data Fig. 9; BRT20220306 and BRTQ20220711) according to the AVMA Guidelines for the Euthanasia of Animals: 2020 edition and in compliance with all applicable state and federal laws.

All experiments involving animals reported in Fig. 5e,f were approved by Washington University in St. Louis IACUC (Animal Welfare Assurance D16-00245). All facilities involving animal research were accredited by the Association for Assessment and Accreditation of Laboratory Animal Care

International. Procedures, facilities and housing were in compliance with the NRC Guide for the Care and Use of Laboratory Animals. Mice were euthanized at 4 days post-infection according to the AVMA Guidelines for the Euthanasia of Animals: 2020 edition and in compliance with all applicable state and federal laws.

A power analysis (80%) with type 1 error set at 5% based on an estimated effect size and historical data were used to determine animal numbers for experiments and to keep the number of animals used to a minimum.

Reporting summary

Further information on research design is available in the Nature Portfolio Reporting Summary linked to this article.

Data availability

The cryo-EM maps and atomic coordinates have been deposited to the Electron Microscopy Data Bank (EMDB) and the PDB with accession numbers: 8G3M and EMD-29703 for 3-FNI9-Fab-bound or 8G3N and EMD-29705 for 4-FNI9-Fab-bound NA from the A/Tanzania/205/2010 strain; 8G3O and EMD-29706 for 3-FNI9-Fab-bound or 8G3P and EMD-29707 for 4-FNI9-Fab-bound NA from the A/Hong_Kong/2671/2019 strain; 8G3Q and EMD-29708 for 3-FNI17-Fab-bound NA from the A/Tanzania/205/2010 strain; 8G3R and EMD-29709 for 1-FNI17-Fab-bound NA from the A/Tanzania/205/2010 strain with S245N and S247T mutations; 8G30 and EMD-29686 for 4-FNI19-Fab-bound NA from the A/Tanzania/205/2010 H3N2 strain; 8G40 and EMD-29712 for 3-FNI19-Fab-bound or 8G3V and EMD-29710 for 4-FNI19-Fab-bound NA from the A/Hong_Kong/2671/2019 strain; and 8G3Z and EMD-29711 for 4-FNI17-Fab-bound NA from the B/Massachusetts/02/2012 (Yamagata) strain. A summary of samples with corresponding EMDB and PDB accession codes can be found in Supplementary Table 5. All datasets generated and information presented in the study are available from the corresponding authors on reasonable request. Materials generated in this study can be available on request and may require a material transfer agreement. Source data are provided with this paper.

References

1. Kirkpatrick, E., Qiu, X., Wilson, P. C., Bahl, J. & Krammer, F. The influenza virus hemagglutinin head evolves faster than the stalk domain. *Sci. Rep.* **8**, 10432 (2018).

2. Paules, C. I., Sullivan, S. G., Subbarao, K. & Fauci, A. S. Chasing seasonal influenza—the need for a universal influenza vaccine. *N. Engl. J. Med.* **378**, 7–9 (2018).

3. GBD 2017 Influenza Collaborators. Mortality, morbidity, and hospitalisations due to influenza lower respiratory tract infections, 2017: an analysis for the Global Burden of Disease Study 2017. *Lancet Respir. Med.* **7**, 69–89 (2019).

4. Subbarao, K. The critical interspecies transmission barrier at the animal–human interface. *Trop. Med. Infect. Dis.* **4**, 72 (2019).

5. Zaraket, H., Hurt, A. C., Clinch, B., Barr, I. & Lee, N. Burden of influenza B virus infection and considerations for clinical management. *Antiviral Res.* **185**, 104970 (2021).

6. Gaymard, A., Le Briand, N., Frobert, E., Lina, B. & Escuret, V. Functional balance between neuraminidase and haemagglutinin in influenza viruses. *Clin. Microbiol. Infect.* **22**, 975–983 (2016).

7. Skehel, J. J. & Wiley, D. C. Receptor binding and membrane fusion in virus entry: the influenza hemagglutinin. *Annu. Rev. Biochem.* **69**, 531–569 (2000).

8. McAuley, J. L., Gilbertson, B. P., Trifkovic, S., Brown, L. E. & McKimm-Breschkin, J. L. Influenza virus neuraminidase structure and functions. *Front. Microbiol.* **10**, 39 (2019).

9. Matrosovich, M. N., Matrosovich, T. Y., Gray, T., Roberts, N. A. & Klenk, H. D. Neuraminidase is important for the initiation of influenza virus infection in human airway epithelium. *J. Virol.* **78**, 12665–12667 (2004).

10. Cohen, M. et al. Influenza A penetrates host mucus by cleaving sialic acids with neuraminidase. *Virol. J.* **10**, 321 (2013).

11. Guo, H. et al. Kinetic analysis of the influenza A virus HA/NA balance reveals contribution of NA to virus-receptor binding and NA-dependent rolling on receptor-containing surfaces. *PLoS Pathog.* **14**, e1007233 (2018).

12. de Vries, E., Du, W., Guo, H. & de Haan, C. A. M. Influenza A virus hemagglutinin-neuraminidase-receptor balance: preserving virus motility. *Trends Microbiol.* **28**, 57–67 (2020).

13. Harris, A. et al. Influenza virus pleiomorphy characterized by cryoelectron tomography. *Proc. Natl Acad. Sci. USA* **103**, 19123–19127 (2006).

14. Monto, A. S. & Kendal, A. P. Effect of neuraminidase antibody on Hong Kong influenza. *Lancet* **1**, 623–625 (1973).

15. Memoli, M. J. et al. Evaluation of antihemagglutinin and antineuraminidase antibodies as correlates of protection in an influenza A/H1N1 virus healthy human challenge model. *mBio* **7**, e00417-16 (2016).

16. Couch, R. B. et al. Antibody correlates and predictors of immunity to naturally occurring influenza in humans and the importance of antibody to the neuraminidase. *J. Infect. Dis.* **207**, 974–981 (2013).

17. Maier, H. E. et al. Pre-existing antineuraminidase antibodies are associated with shortened duration of influenza A(H1N1)pdm virus shedding and illness in naturally infected adults. *Clin. Infect. Dis.* **70**, 2290–2297 (2020).

18. DiLillo, D. J., Palese, P., Wilson, P. C. & Ravetch, J. V. Broadly neutralizing anti-influenza antibodies require Fc receptor engagement for in vivo protection. *J. Clin. Invest.* **126**, 605–610 (2016).

19. Krammer, F. et al. NAction! How can neuraminidase-based immunity contribute to better influenza virus vaccines? *mBio* **9**, e02332-17

(2018).

20. Lim, J. J. et al. A phase 2 randomized, double-blind, placebo-controlled trial of MHAA4549A, a monoclonal antibody, plus oseltamivir in patients hospitalized with severe influenza A virus infection. *Antimicrob. Agents Chemother.* **64**, 1285–1210 (2020).

21. Ali, S. O. et al. Evaluation of MEDI8852, an anti-influenza A monoclonal antibody, in treating acute uncomplicated influenza. *Antimicrob. Agents Chemother.* **62**, 1–47 (2018).

22. US National Library of Medicine. *ClinicalTrials.gov* https://clinicaltrials.gov/ct2/show/NCT05567783 (2022).

23. Yasuhara, A. et al. A broadly protective human monoclonal antibody targeting the sialidase activity of influenza A and B virus neuraminidases. *Nat. Commun.* **13**, 6602 (2022).

24. Stadlbauer, D. et al. Broadly protective human antibodies that target the active site of influenza virus neuraminidase. *Science* **366**, 499–504 (2019).

25. Madsen, A. et al. Human antibodies targeting influenza B virus neuraminidase active site are broadly protective. *Immunity* **53**, 852–863.e7 (2020).

26. Chen, Y. Q. et al. Influenza infection in humans induces broadly cross-reactive and protective neuraminidase-reactive antibodies. *Cell* **173**, 417–429.e10 (2018).

27. Hurt, A. C. et al. Community transmission of oseltamivir-resistant A(H1N1)pdm09 influenza. *N. Engl. J. Med.* **365**, 2541–2542 (2011).

28. Meijer, A. et al. Oseltamivir-resistant influenza virus A (H1N1), Europe, 2007–08 season. *Emerg. Infect. Dis.* **15**, 552–560 (2009).

29. Powell, H. & Pekosz, A. Neuraminidase antigenic drift of H3N2 clade 3c.2a viruses alters virus replication, enzymatic activity and inhibitory antibody binding. *PLoS Pathog.* **16**, e1008411 (2020).

30. Wan, H. et al. The neuraminidase of A(H3N2) influenza viruses circulating since 2016 is antigenically distinct from the A/Hong Kong/4801/2014 vaccine strain. *Nat. Microbiol.* **4**, 2216–2225 (2019).

31. Stadlbauer, D. et al. Antibodies targeting the neuraminidase active site inhibit influenza H3N2 viruses with an S245N glycosylation site. *Nat. Commun.* **13**, 7864 (2022).

32. Ke, C. et al. Human infection with highly pathogenic avian influenza A(H7N9) virus, China. *Emerg. Infect. Dis.* **23**, 1332–1340 (2017).

33. Aguero, M. et al. Highly pathogenic avian influenza A(H5N1) virus infection in farmed minks, Spain, October 2022. *Euro. Surveill.* https://doi.org/10.2807/1560-7917.ES.2023.28.3.2300001 (2023).

34. Kallewaard, N. L. et al. Structure and function analysis of an antibody recognizing all influenza A subtypes. *Cell* **166**, 596–608 (2016).

35. Gamblin, S. J. & Skehel, J. J. Influenza hemagglutinin and neuraminidase membrane glycoproteins. *J. Biol. Chem.* **285**, 28403–28409 (2010).

36. Sazinsky, S. L. et al. Aglycosylated immunoglobulin G1 variants productively engage activating Fc receptors. *Proc. Natl Acad. Sci. USA* **105**, 20167–20172 (2008).

37. DiLillo, D. J., Tan, G. S., Palese, P. & Ravetch, J. V. Broadly neutralizing hemagglutinin stalk–specific antibodies require FcγR interactions for protection against influenza virus in vivo. *Nat. Med.* **20**, 143–151 (2014).

38. Govorkova, E. A. et al. Susceptibility of highly pathogenic H5N1 influenza viruses to the neuraminidase inhibitor oseltamivir differs in vitro and in a mouse model. *Antimicrob. Agents Chemother.* **53**, 3088–3096 (2009).

39. Oakley, A. J. et al. Structural and functional basis of resistance to neuraminidase inhibitors of influenza B viruses. *J. Med. Chem.* **53**, 6421–6431 (2010).

40. Ison, M. G. et al. Early treatment with baloxavir marboxil in high-risk adolescent and adult outpatients with uncomplicated influenza (CAPSTONE-2): a randomised, placebo-controlled, phase 3 trial. *Lancet Infect. Dis.* **20**, 1204–1214 (2020).

41. Henry, C., Palm, A. E., Krammer, F. & Wilson, P. C. From original antigenic sin to the universal influenza virus caccine. *Trends Immunol.* **39**, 70–79 (2018).

42. Hayden, F. G. et al. Use of the selective oral neuraminidase inhibitor oseltamivir to prevent influenza. *N. Engl. J. Med.* **341**, 1336–1343 (1999).

43. Hayden, F. G. et al. Inhaled zanamivir for the prevention of influenza in families. Zanamivir Family Study Group. *N. Engl. J. Med.* **343**, 1282–1289 (2000).

44. Ng, S. et al. Novel correlates of protection against pandemic H1N1 influenza A virus infection. *Nat. Med.* **25**, 962–967 (2019).

45. Cheung, C. S. et al. Identification and structure of a multidonor class of head-directed influenza-neutralizing antibodies reveal the mechanism for its recurrent elicitation. *Cell Rep.* **32**, 108088 (2020).

46. Zalevsky, J. et al. Enhanced antibody half-life improves in vivo activity. *Nat. Biotechnol.* **28**, 157–159 (2010).

47. Leon, P. E. et al. Optimal activation of Fc-mediated effector functions by influenza virus hemagglutinin antibodies requires two points of contact. *Proc. Natl Acad. Sci. USA* **113**, E5944–E5951 (2016).

48. Zhang, A. et al. Hemagglutinin stalk-binding antibodies enhance effectiveness of neuraminidase inhibitors against influenza via Fc-dependent effector functions. *Cell Rep. Med.* **3**, 100718 (2022).

49. Ellis, D. et al. Structure-based design of stabilized recombinant influenza neuraminidase tetramers. *Nat. Commun.* **13**, 1825 (2022).

50. Strohmeier, S. et al. A novel recombinant influenza virus neuraminidase vaccine candidate stabilized by a measles virus phosphoprotein tetramerization domain provides robust protection from virus challenge in the mouse model. *mBio* **12**, e0224121 (2021).

51. Xue, J., Chambers, B. S., Hensley, S. E. & Lopez, C. B. Propagation and characterization of influenza virus stocks that lack high levels of defective viral genomes and hemagglutinin mutations. *Front. Microbiol.* **7**, 326 (2016).

52. Tiller, T. et al. Efficient generation of monoclonal antibodies from single human B cells by single cell RT-PCR and expression vector cloning. *J. Immunol. Methods* **329**, 112–124 (2008).

53. Foglierini, M., Pappas, L., Lanzavecchia, A., Corti, D. & Perez, L. AncesTree: an interactive immunoglobulin lineage tree visualizer. *PLoS Comput. Biol.* **16**, e1007731 (2020).

54. Kozlov, A. M., Darriba, D., Flouri, T., Morel, B. & Stamatakis, A. RAxML-NG: a fast, scalable and user-friendly tool for maximum likelihood phylogenetic inference. *Bioinformatics* **35**, 4453–4455 (2019).

55. van Baalen, C. A. et al. ViroSpot microneutralization assay for antigenic characterization of human influenza viruses. *Vaccine* **35**, 46–52 (2017).

56. Reed, L. J. & Muench, H. A simple method of estimating fifty per cent endpoints. *Am. J. Epidemiol.* **27**, 493–497 (1938).

57. Mastronarde, D. N. Automated electron microscope tomography using robust prediction of specimen movements. *J. Struct. Biol.* **152**, 36–51 (2005).

58. Tan, Y. Z., Cheng, A., Potter, C. S. & Carragher, B. Automated data collection in single particle electron microscopy. *Microscopy* **65**, 43–56 (2016).

59. Scheres, S. H. Processing of structurally heterogeneous cryo-EM data in RELION. *Methods Enzymol.* **579**, 125–157 (2016).

60. Kimanius, D., Forsberg, B. O., Scheres, S. H. & Lindahl, E. Accelerated cryo-EM structure determination with parallelisation using GPUs in RELION-2. *eLife* **5**, e18722 (2016).

61. Zheng, S. Q. et al. MotionCor2: anisotropic correction of beam-induced motion for improved cryo-electron microscopy. *Nat. Methods* **14**, 331–332 (2017).

62. Rohou, A. & Grigorieff, N. CTFFIND4: fast and accurate defocus estimation from electron micrographs. *J. Struct. Biol.* **192**, 216–221 (2015).

63. Punjani, A., Rubinstein, J. L., Fleet, D. J. & Brubaker, M. A. cryoSPARC: algorithms for rapid unsupervised cryo-EM structure determination. *Nat. Methods* **14**, 290–296 (2017).

64. Zhang, K. Gctf: real-time CTF determination and correction. *J. Struct. Biol.* **193**, 1–12 (2016).

65. Scheres, S. H. & Chen, S. Prevention of overfitting in cryo-EM structure determination. *Nat. Methods* **9**, 853–854 (2012).

66. Pettersen, E. F. et al. UCSF Chimera—a visualization system for exploratory research and analysis. *J. Comput. Chem.* **25**, 1605–1612 (2004).

67. Emsley, P. & Cowtan, K. Coot: model-building tools for molecular graphics. *Acta Crystallogr. D Biol. Crystallogr.* **60**, 2126–2132 (2004).

68. Croll, T. I. ISOLDE: a physically realistic environment for model building into low-resolution electron-density maps. *Acta Crystallogr. D Struct. Biol.* **74**, 519–530 (2018).

69. Nicholls, R. A., Fischer, M., McNicholas, S. & Murshudov, G. N. Conformation-independent structural comparison of macromolecules

with ProSMART. *Acta Crystallogr. D Biol. Crystallogr.* **70**, 2487–2499 (2014).

70. Yamashita, K., Palmer, C. M., Burnley, T. & Murshudov, G. N. Cryo-EM single-particle structure refinement and map calculation using Servalcat. *Acta Crystallogr. D Struct. Biol.* **77**, 1282–1291 (2021).

71. Brown, A. et al. Tools for macromolecular model building and refinement into electron cryo-microscopy reconstructions. *Acta Crystallogr. D Biol. Crystallogr.* **71**, 136–153 (2015).

72. Schrödinger, L. D. W. PyMOL. *PyMOL* http://www.pymol.org/pymol (2020).

73. Starr, T. N. et al. SARS-CoV-2 RBD antibodies that maximize breadth and resistance to escape. *Nature* **597**, 97–102 (2021).

74. Case, D. A. et al. Amber: Tools for Molecular Simulations. *University of California, San Francisco.* https://ambermd.org/index.php (2023).

75. Maier, J. A. et al. ff14SB: improving the accuracy of protein side chain and backbone parameters from ff99SB. *J. Chem. Theory Comput.* **11**, 3696–3713 (2015).

76. Kirschner, K. N. et al. GLYCAM06: a generalizable biomolecular force field. Carbohydrates. *J. Comput. Chem.* **29**, 622–655 (2008).

77. Jorgensen, W. L., Chandrasekhar, J., Madura, J. D., Impey, R. W. & Klein, M. L. Comparison of simple potential functions for simulating liquid water. *J. Chem. Phys.* **79**, 926–935 (1983).

78. Joung, I. S. & Cheatham, T. E. 3rd Determination of alkali and halide monovalent ion parameters for use in explicitly solvated biomolecular simulations. *J. Phys. Chem. B* **112**, 9020–9041 (2008).

79. Starr, T. N. et al. Shifting mutational constraints in the SARS-CoV-2 receptor-binding domain during viral evolution. *Science* **377**, 420–424 (2022).

80. Roe, D. R. & Cheatham, T. E. 3rd Parallelization of CPPTRAJ enables large scale analysis of molecular dynamics trajectory data. *J. Comput. Chem.* **39**, 2110–2117 (2018).

81. Yang X. et al. Defining protein variant functions using high-complexity mutagenesis libraries and enhanced mutant detection software ASMv1.0. Preprint at *bioRxiv* https://doi.org/10.1101/2021.06.16.448102 (2023).

82. Katoh, K. & Standley, D. M. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Mol. Biol. Evol.* **30**, 772–780 (2013).

83. Wagih, O. ggseqlogo: A versatile R package for drawing sequence logos. *Bioinformatics* **33**, 3645–3647 (2017).

Acknowledgements

We thank NanoImaging Services for some of the Cryo-EM data collection and processing; S. A. Kristian and R. Ray (GSK) for their critical reading and scientific insights; the NIBSC and IRR for providing seasonal influenza viruses; and X. Yang, D. E. Root and O. Bare at the Genetic Perturbation Platform (GPP) of the Broad Institute of MIT and Harvard for assistance with the deep mutational scanning library construction, sequencing and variant calling.

Author information

Author notes

1. These authors contributed equally: Corey Momont, Ha V. Dang, Fabrizia Zatta, Kevin Hauser

Authors and Affiliations

1. Vir Biotechnology, San Francisco, CA, USA

Corey Momont, Ha V. Dang, Kevin Hauser, Julia di Iulio, Nadine Czudnochowski, Kaitlin Branch, Siddhant Vyas, Alex Chen, Elena Ferri, Abigail E. Powell, Roberto Spreafico, Samantha S. Yim, Dale R. Balce, Istvan Bartha, Marcel Meury, Jessica L. Miller, Amalio Telenti, Herbert W. Virgin, Laura E. Rosen, Lisa A. Purcell & Gyorgy Snell

2. Humabs Biomed SA, a subsidiary of Vir Biotechnology, Bellinzona, Switzerland

   Fabrizia Zatta, Andrea Minola, Anna De Marco, Barbara Guarino, Michael A. Schmid, Elisabetta Cameroni, Antonio Lanzavecchia, Davide Corti & Matteo Samuele Pizzuto

3. Vir Biotechnology, St. Louis, MO, USA

   Caihong Wang, David Donermeyer & William Timothy Schaiff

4. Cambridge Institute for Medical Research, Department of Haematology, University of Cambridge, Cambridge, UK

   Tristan I. Croll

5. School of Biological Sciences, Department of Biochemistry, University of Utah, Salt Lake City, UT, USA

   David M. Belnap

6. Department of Pathology and Immunology, Washington University School of Medicine, St. Louis, MO, USA

   Herbert W. Virgin

Contributions

G.S., D.C. and M.S.P. conceived the research and designed the study. C.M., H.V.D., F.Z., K.H., C.W., J.d.I., A.D.M., K.B., D.D., B.G., L.E.R, L.A.P., G.S., D.C. and M.S.P. came up with the experimental design. F.Z., A.M., A.D.M. and E.C. took part in donor recruitment and PBMC sample

processing. F.Z., A.M., A.D.M. and E.C. conducted mAb isolation, recombinant expression and purification. C.M., F.Z., A.D.M., K.B. and S.V. participated in binding, MUNANA, ELLA and neutralization assays. F.Z. and M.S.P. designed the expression plasmids for transient expression in mammalian cells and pseudoparticle production. H.V.D., N.C., E.F. and G.S. undertook NA antigen design and production. H.V.D. and L.E.R. did the affinity measurements by SPR assays. E.F., A.E.P. and N.C. conducted complex formation for cryo-EM. H.V.D., K.H., E.F., A.C., T.I.C. and D.M.B. performed cryo-EM data collection, processing and model building. A.D.M., B.G. and M.A.S. evaluated effector functions. A.D.M., K.B., S.S.Y., D.R.B., I.B. and M.M. performed resistance studies and deep mutational scanning. J.d.I., R.S., I.B. and A.T. conducted the bioinformatic epitope conservation and clonal evolution analysis. C.W., D.D. and W.T.S. contributed the BALB/c model and performed the in vivo studies. C.M., H.V.D., F.Z., K.H., C.W., J.d.I., N.C., A.D.M., K.B., D.D., B.G., A.E.P., L.E.R. and M.S.P. performed data analysis. J.L.M. and M.S.P. managed the project. C.M., N.C., M.A.S., W.T.S., J.L.M., E.C., A.T., H.W.V., L.E.R., L.A.P., G.S., D.C. and M.S.P. supervised the project. A.L., G.S., D.C. and M.S.P. wrote the manuscript with input from all the authors.

Corresponding authors

Correspondence to Gyorgy Snell, Davide Corti or Matteo Samuele Pizzuto.

Ethics declarations

Competing interests

C.M., H.V.D., F.Z., K.H., C.W., J.d.I., A.M., N.C., A.D.M., K.B., D.D., S.V., A.C., B.G., A.E.P., S.S.Y., D.R.B., I.B., M.M., M.A.S., W.T.S., J.L.M., E.C., A.T., L.E.R., L.A.P., A.L., G.S., D.C. and M.S.P. are employees of and may hold shares in Vir Biotechnology. D.C., M.S.P., A.M., E.C., G.S., K.H., C.M. and E.F. are currently listed as inventors on multiple patent applications, which disclose the subject matter described in this paper. E.F., A.E.P., R.S. and H.W.V. are former employees of VIR Biotechnology. E.F. is currently an employee at and hold stocks in Amgen. Amgen provided no funding for this work. T.I.C. is a consultant for Vir Biotechnology and

receives royalties from licensing of ISOLDE software used for molecular modelling. D.M.B. received research funding from VIR Biotechnology. L.A.P. is a former employee and shareholder of Regeneron Pharmaceuticals and is a member of the Scientific Advisory Board AI-driven structure-enabled antiviral platform (ASAP). Regeneron provided no funding for this work. H.W.V. is a founder of PierianDx and Casma Therapeutics. Neither company provided funding for this work. The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential competing interest.

Peer review

Peer review information

Nature thanks the anonymous reviewers for their contribution to the peer review of this work.

Additional information

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Extended data figures and tables

Extended Data Fig. 1 Clonal evolution of FNI mAbs.

a, b, Graphical representation of the FNI clonal evolution based on the VH (a) and VK (b) generated by the graphic user interface (GUI) AncesTree[53]. Branch lengths are denoted on the lines connecting the branching points with both the total number of nucleotide and amino acid changes, the latter in parentheses. **c**, Phylogenetic tree of the concatenated heavy and light chain sequences generated using raxml-ng[54] assuming the same single somatic substitution process across both heavy and light chain genes. Each heavy chain sequence was appended to the light chain sequence of the same clone for the purpose of phylogenetic reconstruction under the assumption of

a single substitution process. Branch points represent hypothetical bifurcations during the evolution of the antibody sequences. Scale represents the number of amino acid changes per site. **d, e**, Amino acid sequence alignment of FNI clonal family VH (d) and VK (e) with respective unmutated common ancestor (UCA).

Extended Data Fig. 2 Breadth of FNI mAbs across the Influenza A and B viruses.

a, FACS gating strategy used to assess binding of FNI and 1G01 mAbs to NA transiently expressed in mammalian cells. **b**, Inhibition of NA enzymatic activity by the anti-NA mAbs against group 1 (N1), group 2 (N2) IAV and B/Victoria/2/87-like (Vic) and B/Yamagata/16/88-like (Yam) IBV NA antigens as measured by MUNANA assay. Data represent n = 1 biologically independent experiment out of two. **c**, Introduction and spreading of the 245 glycosylation motif in H3N2 seasonal viruses. GISAID analysis (2011–2022) of the N2 NAs from seasonal H3N2 strains bearing S245N and S247T mutations, which results in the introduction of the N245 glycosylation site. X represents strains selected for inclusion in the vaccine formulations. Source: Nextstrain (https://nextstrain.org/flu/seasonal/h3n2/na/12y?c=gt-NA_245,247&dmin=2004-10-11).

Source Data

Extended Data Fig. 3 NAI activity of FNI and 1G01 mAbs against pseudoparticles bearing NAs from avian and mammalian influenza A viruses.

a, Inhibition of sialidase activity of FNI and 1G01 mAbs versus NA-based pseudoparticles bearing N1 from highly pathogenic H5N1 A/Mink/Spain/3691-8_22VIR10586-10/2022 (clade 2.3.4.4) as measure by MUNANA assay. **b**, Inhibition of neuraminidase enzymatic activity, as measured by ELLA, exerted by anti-NA mAbs against NA-based pseudoparticles bearing N3, N4, or N5 representative of enzootic low pathogenic avian influenza A viruses. **c**, Inhibition of sialidase activity of FNI and 1G01 mAbs versus NA-only based pseudoparticles bearing NA representative of enzootic swine or canine influenza A viruses as measured

by MUNANA assay. Data from n = 1 biologically independent experiment out of two are shown.

Source Data

Extended Data Fig. 4 FNI9 mAb mediates activation of CDC, ADCC and ADCP.

a,d, CDC with a serial dilution of FNI9 (red) and anti-HA MEDI8852 (green) mAbs on MDCK-LN cells infected with H1N1 PR8 in the presence of guinea pig complement. **b,e**, ADCC with FNI9 (red) and anti-HA MEDI8852 (green) mAbs on A549 cells infected with H1N1 PR8 in the presence of freshly isolated human NK cells. **c,f**, ADCP with serial dilution of FNI9 mAb (red) using peripheral blood mononuclear cells (PBMCs as source of monocytes) as effector cells, and PKH67-labelled ExpiCHO cells expressing N2 NA as target cells. The y-axis indicates the percentage of monocytes double positive for CD14 and PKH67. For all assays FNI9-GRLR (red empty circle) is used as Fc-silent negative control and results are shown both as dose-response curves (a-c) and as area under the curve (AUC) (d-f). Results are representative of n = 1 (d, e) or n = 2 (f) biological replicates (black dots).

Source Data

Extended Data Fig. 5 Conservation analysis of FNI mAbs key contact residues on the NA of IAVs and IBVs.

Mean conservation percentage (2009–2022) of key NA contacting residues (R118, D151, E227, R292, and R371) per year (red line). The number of sequences analysed per year is shown (black bars). Sequences retrieved for H1N1, H3N2 and IBV isolates are of human origin, while those for H5N1, H7N9, H5N8 and H5N6 viruses are of both human and animal origin. Sequences were retrieved from GISAID database. IAV sequences conservation was measured against A/New York/392/2004 reference strain. IBV sequences conservation was measured against B/Yamagata/16/88 reference strain.

Source Data

Extended Data Fig. 6 Escape mutants, frequency and fitness cost.

a, Summary table showing amino acid substitutions identified in resistance studies with H1N1 and H3N2 viruses, the number of virus passages that were required to observe the mutation, the neutralization IC50s against the mutants and the frequency of the mutation in all IAVs deposited in GISAID between 2000–2022. NT*: not tested as R152I and D199N mutants were not rescued following virus propagation in vitro. Note: IC50s observed with S247R/I223R and K431E mutants are in the range of the overall IC50s observed for FNI9 with seasonal IAVs and IBVs (see Fig. 2a). **b**, Mutations reducing FNI9 binding to N1 from H1N1 A/California/07/2009 identified by DMS. Only amino acid substitutions that do not result in strong reduction of expression are reported in the red colour scale (darker red, more expression; lighter red, less expression in comparison to synonymous mutations). N1 numbering is used on the x-axis along with N2 numbering below. N2 positions identified as epitope residues by static structural analysis are highlighted with black dots (‘Epi’) and additional positions identified by dynamic analysis (MD) are highlighted with red dots. Potential escape mutations flagged by DMS (top row of black dots, ‘Esc’) not identified as epitope residues have low frequencies in GISAID (see below). **c**, Static epitope analysis of SA, oseltamivir (OSE), FNI9, FNI17, FNI19, and 1G01 shown as in Fig. 3h, where lighter red is a weaker interaction and darker red is a stronger interaction. **d**, Logo plot of N1 sequences in GISAID shows the frequencies of all residues at each position across all species (n = 79,316). N1 numbering is used on the x-axis and the y-axis extends from a cumulative frequency of zero to one. The potential escape mutants in (b) are too infrequent to be seen in (d). **e**, Frequency of only the DMS-identified potential escapes appearing in GISAID for N1 across all species and all time windows (n = 79,316) shown as a logo plot. N1 numbering is used on the x-axis and the y-axis extends from a cumulative frequency of zero to 10^{-3}. **f**, Enzymatic activity, as measured by MUNANA, of N1 antigen from H5N1 A/Vietnam/1203/2004 or N2 antigen from H3N2 A/Hong Kong/2671/2019 bearing amino acid substitutions identified via resistance studies (a) in comparison to the sialidase activity of the wild type NAs. A detailed list of

the mutations and NA antigens in which they were introduced is reported in Supplemetary Table 4. **g**, Enzymatic activity, as measured by MUNANA, of pseudoparticles bearing N1 from H1N1 A/California/07/09 with amino acid substitutions retrieved from DMS (b-e) in comparison to pseudoparticles presenting the parental NA. **h,i**, Inhibition of neuraminidase enzymatic activity by FNI and 1G01 mAbs, as measured by MUNANA, of a panel of NA bearing mutations identified through resistance studies (h) or DMS (i). Data for n = 1 biologically independent experiment are shown.

Source Data

Extended Data Fig. 7 Glycan profiling of neuraminidase with peptide mapping LC-MS.

a, Glycan profile of N2 A/Tanzania/2010 with pie charts depicting the percent distributions of glycans. The most abundant glycan (red font) is illustrated; the monosaccharide units resolved by cryoEM are outlined in red. The most abundant glycan at each position in N2 was modelled onto cryoEM structures to prepare the MD simulations (see Methods for details). **b**, Glycan profile of N2 A/Hong Kong/2019 depicted as in (a). Oxford glycan notation is used; nG, no glycan.

Source Data

Extended Data Fig. 8 Detailed structural analysis of FNI9 and FNI17 mAbs binding to NA.

a, Static epitope analysis of FNI9:NA (left) and FNI17:NA (right) on N2 A/Tanzania/2010 depicted as a heatmap whose color scale is MOE kcal/mol energy; x-axis delineates epitope residues; and y-axis delineates paratope residues. FNI9 and FNI17 sequence differences are specified by grey dots along the y-axes and total binding energies are reported. **b**, Dynamic epitope analysis of FNI9:NA (left) and FNI17:NA (right) on N2 A/Tanzania/2010 depicted as in (a) using energies derived from 11.5 μs MD simulations. Red squares depict epitope-paratope residue contacts present in the FNI9:NA complex that are absent in the FNI17:NA complex. Dynamic total binding energies are reported. **c**, Percent occupancy of epitope-paratope interactions

(within 5 Å) of FNI9:NA (left) and FNI17:NA (right) on N2 A/Tanzania/2010 in the 11.5 µs MD simulations. Percentage of total occupancy is reported. **d**, Comparison of FNI9 and FNI17 paratopes using dynamic analysis; FNI9 and FNI17 sequence differences are specified by grey dots along the x-axis. **e**, Comparison of FNI9 and FNI17 epitopes using dynamic analysis; NA residues depicted along the x-axis.

Source Data

Extended Data Fig. 9 Contribution of effector functions and combination studies in vivo.

a, Body weight loss of BALB/c mice (n = 6 mice/group) related to Fig. 5a–d. Body weight loss percentage for single BALB/c mice prophylactically administered with muFNI9, muFNI9 N297Q, muMEDI8852, and muMEDI8852 N297Q mAbs at 9, 3, 0.9, and 0.3 mg/kg and infected with a lethal dose of H1N1 A/Puerto Rico/8/34. **b**, Area of negative peaks of the body weigh loss measured for the BALB/c mice (n = 6 mice/group) prophylactically administered with murinized MEDI8852 (muMEDI8852), murinized FNI9 (muFNI9), MEDI8852, FNI17, FNI9 and the combination of the anti-HA and anti-NA mAbs in a 1:1 ratio against H1N1 A/Puerto Rico/8/34 or H3N2 A/Hong Kong/1/1968 as indicated in the graphs. The area of the negative peaks is defined as the area between the body weight loss line of each animal over 14 days and the 0% body weight loss baseline. Data from n = 1 independent experiment for each anti-NA mAb are presented. Two-tailed Mann–Whitney test was used for statistical analysis of significance. Errors bars indicate S.D.

Source Data

Extended Data Fig. 10 Correlation between mAb concentration at Day 0 and virus lung titres related to Fig. 5e–h.

Replicating virus titres in the lungs of BALB/c mice at 4 days after challenge plotted against serum mAb concentrations measured 2 h before infection with H1N1 A/Puerto Rico/8/34, H3N2 A/Singapore/2016, B/Victoria/504/2000, or B/Brisbane/60/2008 viruses. Dotted red lines

represent viral lung titre reductions produced by oseltamivir. Details on the mAb quantification are reported in the corresponding material and method section.

Source Data

Supplementary information

Reporting Summary

Supplementary Table 1

HCDR3 and LCDR3 sequences and percentage of homology to UCA of clonally related FNI mAbs. This table summarizes the percentage of homology of the FNI mAbs VH and VL regions to unmutated common ancestor (UCA) and reports for each mAb the amino acid sequences of the heavy chain and light chain complementarity-determining regions (HCDR3 and LCDR3).

Supplementary Table 2

List of NA constructs transiently expressed in mammalian cells assessed for FNI9, FNI17, and FNI19 mAbs binding by FACS. This table reports the subtypes, strains, GISAID and/or GenBank accession numbers and relevant references for the NA antigens tested by FACS with FNI9, FNI17, and FNI19 mAbs as shown in Fig.1b.

Supplementary Table 3

List of the IAV/IBV strains and corresponding neutralization IC_{50} values for the anti-NA mAbs related to Fig. 2a. This table shows the full designation of the strains listed as abbreviations in Fig.2a and the corresponding neutralization IC50 values for FNI9, FNI17, FNI19, and 1G01 mAbs.

Supplementary Table 4

List of NA antigens used in the study and corresponding binding affinity values (KD) for FNI9, FNI17, FNI19 and 1G01 Fabs as measured by SPR. This table summarises the binding affinity measure by surface plasmon resonance (SPR) for FNI9, FNI17, FNI19, and 1G01 Fabs versus the N1 or N2 antigens described in the article.

Supplementary Table 5

Cryo-EM data collection, refinement and validation statistics This table summarises the data collection, refinement and validation statistics for the Cryo-EM structures presented in the article.

Source data

Source Data Fig. 2

Source Data Fig. 3

Source Data Fig. 4

Source Data Fig. 5

Source Data Extended Data Fig. 2

Source Data Extended Data Fig. 3

Source Data Extended Data Fig. 4

Source Data Extended Data Fig. 5

Source Data Extended Data Fig. 6

Source Data Extended Data Fig. 7

Source Data Extended Data Fig. 8

[Source Data Extended Data Fig. 9]

[Source Data Extended Data Fig. 10]

Rights and permissions

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by/4.0/.

Reprints and Permissions

About this article

Cite this article

Momont, C., Dang, H.V., Zatta, F. *et al.* A pan-influenza antibody inhibiting neuraminidase via receptor mimicry. *Nature* **618**, 590–597 (2023). https://doi.org/10.1038/s41586-023-06136-y

- Received: 19 September 2022
- Accepted: 26 April 2023
- Published: 31 May 2023
- Issue Date: 15 June 2023

- DOI: https://doi.org/10.1038/s41586-023-06136-y

Share this article

Anyone you share the following link with will be able to read this content:

Get shareable link

Sorry, a shareable link is not currently available for this article.

Copy to clipboard

Provided by the Springer Nature SharedIt content-sharing initiative

---

This article was downloaded by **calibre** from https://www.nature.com/articles/s41586-023-06136-y

- Article
- Open Access
- Published: 31 May 2023

Axonemal structures reveal mechanoregulatory and disease mechanisms

- Travis Walton,
- Miao Gui,
- Simona Velkova,
- Mahmoud R. Fassad,
- Robert A. Hirst,
- Eric Haarman,
- Christopher O'Callaghan,
- Mathieu Bottier,
- Thomas Burgoyne,
- Hannah M. Mitchison &
- …
- Alan Brown

Show authors

Nature **volume 618**, pages 625–633 (2023)

- 11k Accesses
- 225 Altmetric
- Metrics details

Abstract

Motile cilia and flagella beat rhythmically on the surface of cells to power the flow of fluid and to enable spermatozoa and unicellular eukaryotes to swim. In humans, defective ciliary motility can lead to male infertility and a congenital disorder called primary ciliary dyskinesia (PCD), in which impaired clearance of mucus by the cilia causes chronic respiratory infections[1]. Ciliary movement is generated by the axoneme, a molecular machine consisting of microtubules, ATP-powered dynein motors and regulatory complexes[2]. The size and complexity of the axoneme has so far prevented the development of an atomic model, hindering efforts to understand how it functions. Here we capitalize on recent developments in artificial intelligence-enabled structure prediction and cryo-electron microscopy (cryo-EM) to determine the structure of the 96-nm modular repeats of axonemes from the flagella of the alga *Chlamydomonas reinhardtii* and human respiratory cilia. Our atomic models provide insights into the conservation and specialization of axonemes, the interconnectivity between dyneins and their regulators, and the mechanisms that maintain axonemal periodicity. Correlated conformational changes in mechanoregulatory complexes with their associated axonemal dynein motors provide a mechanism for the long-hypothesized mechanotransduction pathway to regulate ciliary motility. Structures of respiratory-cilia doublet microtubules from four individuals with PCD reveal how the loss of individual docking factors can selectively eradicate periodically repeating structures.

Main

The prototypic axoneme of a motile cilium consists of a cylindrical arrangement of doublet microtubules (DMTs) surrounding a central apparatus comprising two singlet microtubules. DMTs are organized into 96-nm repeating units, which cryo-electron tomography (cryo-ET) studies have shown to be a landscape of interconnected dynein motors and regulatory complexes[2]. In total, there are 11 dynein motors in a single 96-nm repeat unit, and they are classified according to their location (inner and outer) and the number of heavy chains (HCs) they have. Each HC contains an ATPase 'head' domain and a helical 'tail' domain that associates with intermediate-chain (IC) and light-chain (LC) subunits. In a 96-nm repeat, there are six different single-headed inner dynein arms (IDA*a*–IDA*e* and IDA*g*), one

double-headed inner dynein arm (IDA*f*) and four identical multi-headed outer dynein arms (ODAs). These ODAs, numbered 1–4 in the proximal to distal direction[3], are organized with 24-nm periodicity in a linear array with extensive head-to-tail interactions[4,5,6]. During ciliary beating, axonemal dyneins docked onto one DMT interact and pull against a second, neighbouring DMT through their motor domains. ODAs provide most of the pulling power, and the six single-headed IDAs are thought to provide torque between DMT pairs[7,8,9]. Axonemal dynein activity switches between specific DMTs on opposite sides of the axoneme[10] through a hypothesized mechanoregulatory pathway involving the central apparatus, DMT-bound T-shaped radial spokes (RSs), the nexin–dynein regulatory complex (N-DRC) and IDA*f*[11].

Despite extensive genetic and tomographic data, the detailed structures of axonemes, and hence their interactions and mechanoregulatory mechanisms, remain poorly understood. We used single-particle cryo-electron microscopy (cryo-EM) and integrative modelling to build atomic models of the 96-nm modular repeats of axonemes from *Chlamydomonas reinhardtii* flagella and human respiratory cilia.

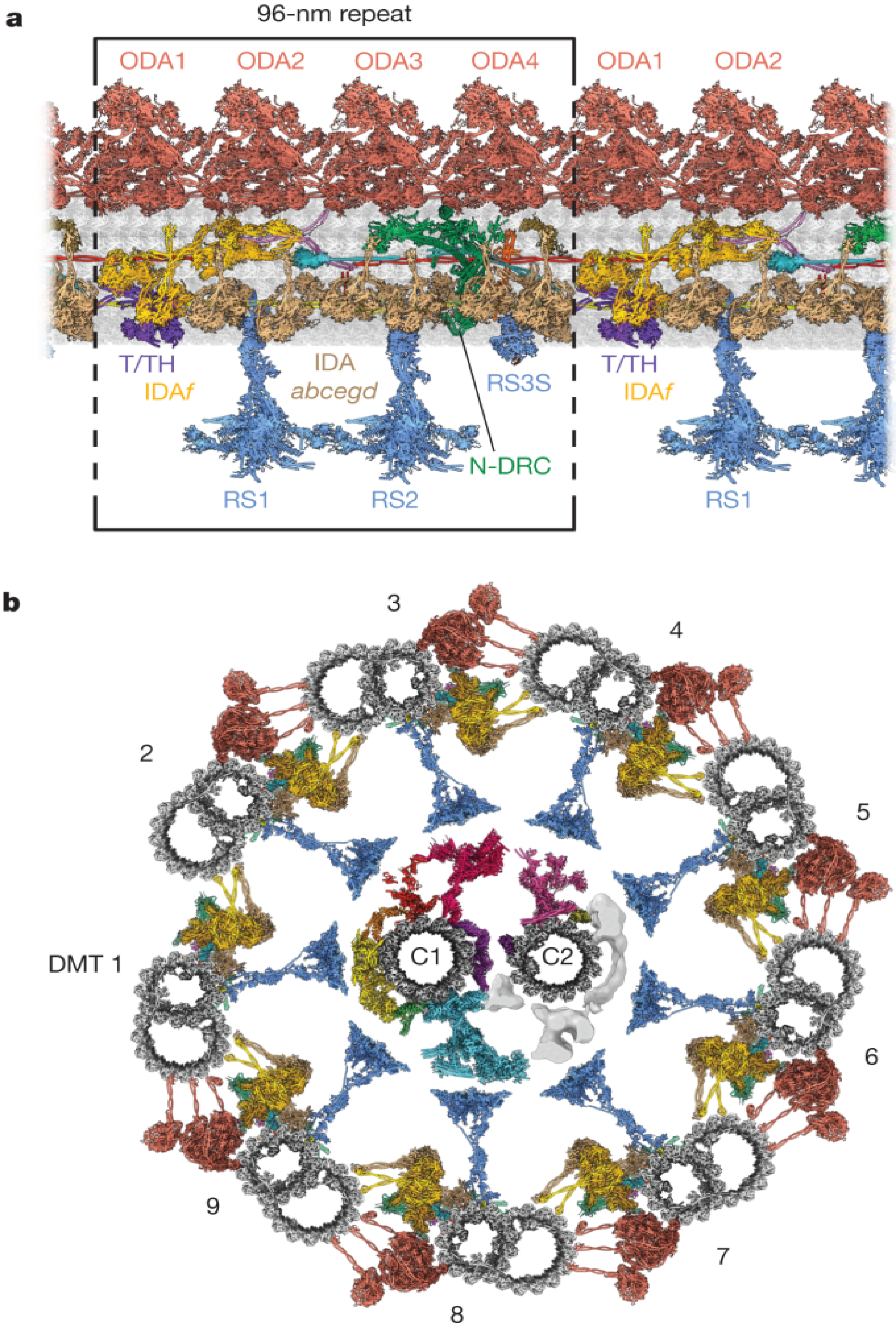
The *C. reinhardtii* axoneme

To determine the structure of the modular repeat of the *C. reinhardtii* axoneme, we reconstructed cryo-EM maps of DMTs from splayed axonemes (Methods and Supplementary Fig. 1). Computational limits imposed by the large size of the repeat meant that we focused on individual axonemal complexes by particle subtraction, refinement and classification before reconstituting the full 96-nm repeat in silico. For some regions, this approach led to structures with resolutions better than 4 Å, whereas other regions had lower resolution because of flexibility that impaired the ability to successfully align the particles, loss of molecular complexes during sample preparation and averaging of asymmetric features (Supplementary Figs. 1–9). The resolution of these maps exceeds what has been achieved previously (about 25 Å) by cryo-ET of *C. reinhardtii* axonemes[12].

To interpret the maps and create atomic models, we used an integrative approach that combined de novo modelling, artificial intelligence (AI)-based

structure prediction (Supplementary Figs. 10–12), cryo-ET studies of *C. reinhardtii* mutants and previous cryo-EM-derived models of doublet microtubule inner proteins (MIPs)[13], RSs[14] and the ODA[5] (Methods). Together with a structure of the central apparatus[15], we generate the most complete atomic model of an axoneme thus far (Fig. 1 and Supplementary Video 1). For all complexes, the subunits identified are provided in Supplementary Table 1 and the rationale for their placement is in Supplementary Table 2. Model statistics are provided in Extended Data Table 1.

Fig. 1: Atomic model of the modular repeat of the *C. reinhardtii* axoneme.

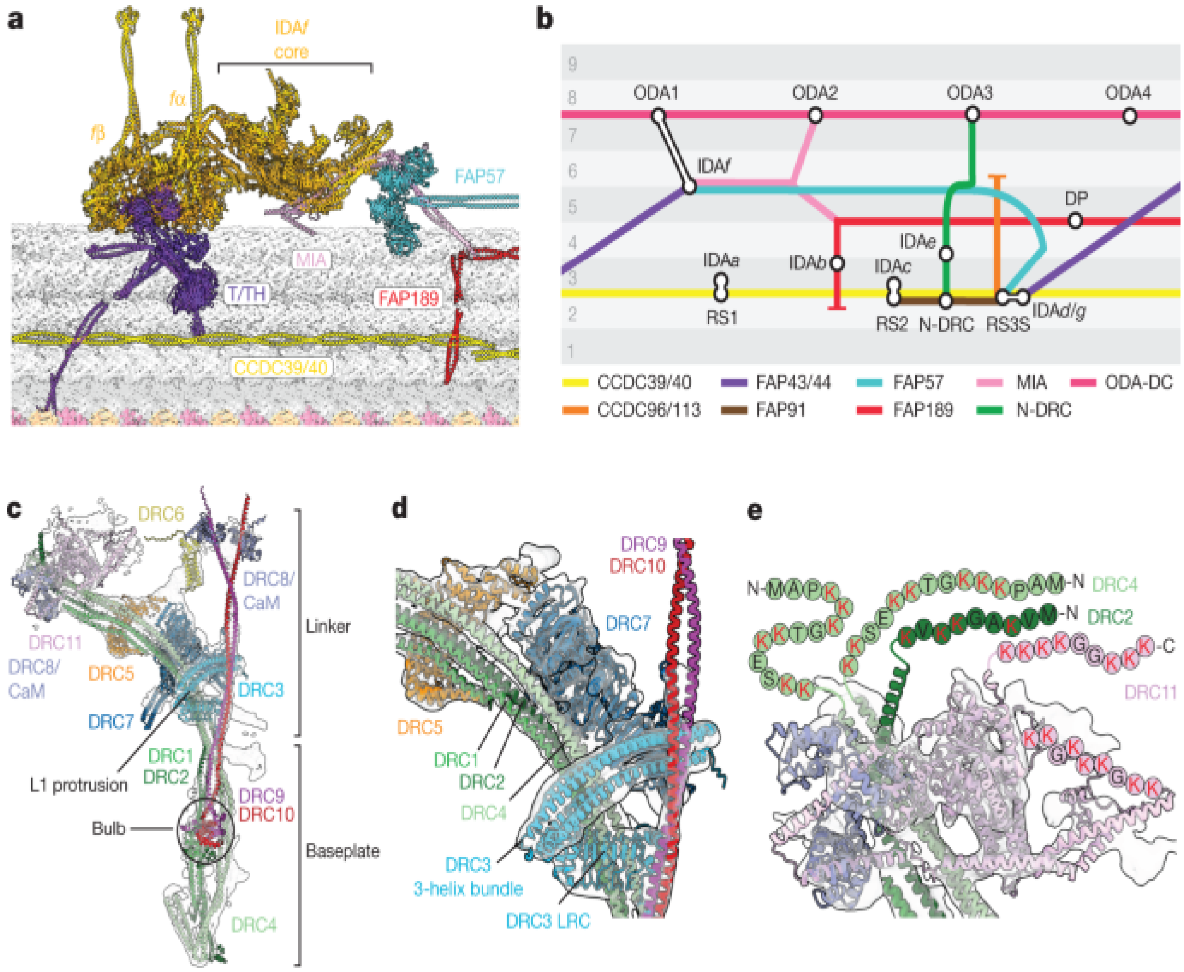

a
96-nm repeat
ODA1
ODA2
ODA3
ODA4
ODA1
ODA2
T/TH
IDAf
IDA
abcegd
RS3S
T/TH
IDAf
N-DRC
RS1
RS2
RS1
b
3
4
2
5
DMT 1
C1
C2
6
9
7
8

a, A longitudinal section of an atomic model of a fully decorated DMT showing the 96-nm repeat boxed. **b**, Cross-sectional model of the *C. reinhardtii* axoneme composed of nine copies of the 96-nm repeat revealed by our study and the C1 and C2 microtubules of the central apparatus (PDB: 7SQC and 7SOM)[15]. The model combines data from multiple sources and provides a consensus view that does not fully capture the asymmetries observed in vivo.

Structures of axonemal dyneins

The model includes all three classes of axonemal dynein: the triple-headed ODA, the regulatory double-headed IDA*f* and all six single-headed IDAs (IDA*a*–IDA*e* and IDA*g*). Although we previously reported[5] a structure of the ODA, the improved accuracy of AI-enabled structure prediction and an atomic model[6] of the ODA from the unicellular ciliate *Tetrahymena thermophila* allowed us to identify two thioredoxin-like subunits (LC3 and LC5) and a second, algal-specific Kelch-like β-propeller domain in the α-HC tail (Extended Data Fig. 1a). We also built an atomic model of IDA*f* (Fig. 2a), which has two ATPase motor domains (*f*α and *f*β) but displays no or little motility in vitro[16,17,18]. Despite their differences in activity, our structure demonstrates that the IDA*f* core (defined as the tails of the HCs and the ICs and LCs with which they associate) closely resembles the ODA core[5] (Extended Data Fig. 1b). Both have an IC–LC block containing two β-propeller-containing ICs, a heterodimer of LC7a–LC7b, three LC8-like dimers and a heterodimer of the TCTEX family. Two IDA*f*-specific subunits, IC97 (ref. [19]) and FAP120 (ref. [20]), expand this region compared with the ODA (Extended Data Fig. 1b). Relative to the core, the motor domains of IDA*f* are rotated and displaced compared with those of the ODA (Extended Data Fig. 1c), probably allowing IDA*f* to bind the neighbouring DMT from its unique position in the 96-nm repeat.

Fig. 2: Structures of mechanoregulatory complexes from the *C. reinhardtii* axoneme.



a, Atomic model depicting the multipartite tethering of IDA*f* by the MIA complex (FAP73/100), a FAP57 homodimer and T/TH. Coiled coils of CCDC39/40 and FAP189 interact with T/TH and MIA, respectively. **b**, A conceptual schematic, in the style of a London Underground map, showing the interconnectivity between axonemal complexes in a 96-nm repeat. Circles represent axonemal complexes, coloured lines represent the path of docking factors and alternating grey zones represent protofilaments. DP, distal protrusion. **c**, Composite cryo-EM map and model of the N-DRC. CaM, calmodulin. **d**, The position of DRC3 at a fork in the N-DRC structure. **e**, DRC2, DRC4 and DRC11 position flexible, lysine-rich regions towards the neighbouring DMT. Individual letters indicate amino acids.

Multipartite docking of IDA*f*

IDA*f* is attached to the DMT through a multipartite docking mechanism involving four different anchors (Fig. 2a,b). The first anchor, which links to ODA1, will be described in detail later (Outer-inner dynein links section). The second anchor corresponds to Modifier of inner arms (MIA), a heterodimeric coiled coil of FAP73 and FAP100 (FAP73/100; ref. [21]) with two separate attachments to the DMT surface (Extended Data Fig. 2a,b). MIA bisects the HC tails of IDA*f* and therefore resembles the DC1–DC2 (DC1/2) coiled coil of the ODA docking complex (ODA-DC), which bisects the γ- and β-HCs of the ODA[5] (Extended Data Fig. 2c). Because IDA*f* has multiple attachment points, MIA is not essential for docking IDA*f* to DMTs[21], but DC1/2 is required for the docking of ODAs[22]. The third anchor of IDA*f* corresponds to a dimer of FAP57 or a member of the FAP57 family[23]. The FAP57 dimer forms a stack of four β-propellers that contacts the *f*α tail (Fig. 2a and Extended Data Fig. 2d), consistent with the proposed location of FAP57 based on subtomogram averaging[23]. The fourth and final anchor is the tether–tetherhead (T/TH) complex[24,25], a stellate structure formed by the entwining of two structurally similar proteins, FAP43 and FAP44 (Extended Data Fig. 3a,b). Dimerization of FAP43 and FAP44 explains why genetic ablation of either factor leads to the loss of the other from mutant cilia[26]. MOT7, a protein reported to regulate *C. reinhardtii* IDA*f* activity in response to blue light[12], occupies the centre of the stellate structure (Extended Data Fig. 3c,d). The N terminus of FAP44 binds the *f*α motor through a pair of β-propellers (Extended Data Fig. 3a) and its C terminus forms a coiled coil with FAP43 across the DMT surface, terminating at the inner junction between the A and B tubules (Extended Data Fig. 3e). The T/TH-mediated connection between IDA*f* and the DMT probably constrains IDA*f* mobility during the powerstroke cycle[24] and undergoes a conformational change in actively beating cilia[10].

N-DRC forms electrostatic DMT links

Our atomic model shows that the 11-subunit N-DRC is built around long, bifurcated coiled coils that are joined together at a ‘bulb’ on the microtubule-bound baseplate and fork at the N-terminal leucine-rich repeat (LRR) domain of DRC3 (Fig. 2c,d and Extended Data Fig. 4a,b). After the fork, coiled coils of DRC1–DRC2 and DRC4–DRC4 curve proximally,

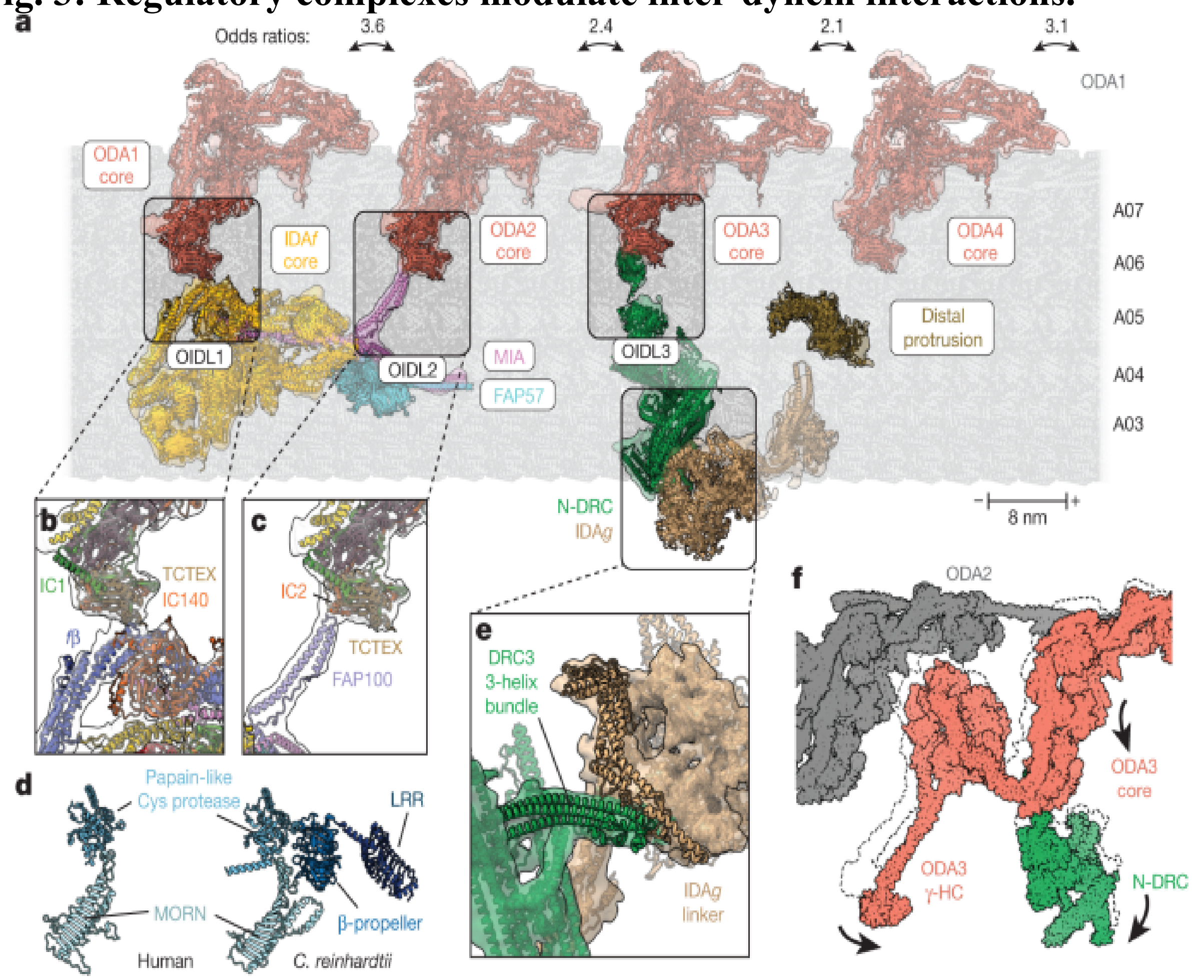
providing sequential binding sites for DRC7, DRC5 and DRC11. These positions agree with subtomogram averages of *C. reinhardtii* mutant axonemes[27] lacking DRC7 and DRC11 (Extended Data Fig. 4c–e) and structural-labelling approaches[28] that positioned DRC1, DRC2, DRC4 and DRC5. The elongated architecture orients six N-DRC subunits with lysine-rich regions in either flexible loops or termini at the interface with the neighbouring DMT. Three of these subunits, DRC2, DRC4 and DRC11, form a cluster on the proximal tip of the N-DRC that contains at least 60 lysine residues (Fig. 2e). On the distal tip, the termini of DRC6, DRC9 and DRC10 point an additional 26 lysine residues towards the neighbouring DMT. These positively charged side chains probably interact electrostatically with negatively charged polyglutamylate post-translational modifications of the neighbouring B-tubule protofilaments[29] to form a physical link with its neighbouring DMT[30]. Molecules of calmodulin (CaM) and/or the calmodulin-like DRC8 subunit bound to IQ motifs on DRC9 to DRC11 near the interface (Fig. 2c,e) suggest that calcium may also regulate the interaction between neighbouring DMTs.

Outer–inner dynein links (OIDLs)

As well as making contact with their neighbouring DMT, both IDA*f* and the N-DRC interact with ODAs and IDAs, forming multiple OIDLs that have a proposed regulatory function to coordinate axonemal dynein activity across the 96-nm repeat unit. IDA*f* contacts both ODA1 and ODA2, and N-DRC contacts ODA3 (ref. [3]; Fig. 3a). The OIDL in each case is mediated by the TCTEX heterodimer (LC2–LC7) of the ODA core, implicating this region as a critical regulatory contact point. ODA1 connects to IDA*f* through the β-propeller domain of IC140 and nearby helical bundles of *f*β (Fig. 3b). This connection has been observed in all axonemes studied thus far, including those with double-headed ODAs[10,31,32]. The second OIDL involves a helical segment of FAP100 that branches off MIA (Fig. 3c). Without MIA, *C. reinhardtii* mutants swim slowly with a low flagellar beat frequency[21,33]. Whether lower motility is due to reduced IDA*f* anchoring or the loss of OIDL2 remains to be investigated. OIDL3 involves the LRR domain of the N-DRC subunit DRC7 (Fig. 3d). Mammalian homologues of DRC7 lack the LRR domain (Fig. 3d), suggesting that this OIDL is specific to algae. We did

not observe a connection between ODA4 and the adjacent distal protrusion[3], a 155-Å-tall armadillo repeat protein that we identify as FAP78 (Extended Data Fig. 5a–d).

Fig. 3: Regulatory complexes modulate inter-dynein interactions.



a, Selected atomic models and their cryo-EM densities showing the longitudinal arrangement of ODA1–ODA4 and their adjacent complexes on the axoneme. Each ODA exists in open and closed conformations, as revealed by 3D classification of the cryo-EM dataset. Odds ratios for the co-occurrence of open and closed states between ODA pairs are shown above the bidirectional arrows. Odds ratios were calculated using 160,444 particles for ODA1, 162,179 particles for ODA2, 214,203 particles for ODA3 and 177,762 particles for ODA4. Scale bar, 8 nm (the length of a tubulin heterodimer) with microtubule polarity indicated. **b**, OIDL1 between the TCTEX of ODA1 and IDA*f* subunits IC140 and *f*β. **c**, OIDL2 between the TCTEX of ODA2 and MIA subunit FAP100. **d**, Models of human (left) and

C. reinhardtii (right) DRC7 with the domains coloured in different shades of blue. Human DRC7 lacks the β-propeller and LRR domains that mediate OIDL3 in *C. reinhardtii*. MORN, membrane occupation and recognition nexus. **e**, Interaction between the three-helix bundle of DRC3 and the linker of IDA*g*. **f**, Overlay of the two major conformational states of N-DRC. Arrows denote movement from the closed to the open state.

We also observed that the N-DRC specifically contacts IDA*g* through DRC3, consistent with earlier subtomogram averages of *drc3* mutant axonemes[34]. Specifically, a three-helix bundle of DRC3 contacts the linker of the IDA*g* motor (Fig. 3e). The linker of dynein motors bends during the powerstroke cycle, suggesting that the interaction between DRC3 and IDA*g* may regulate dynein activity. Alternatively, this interaction may also help position or rotate the IDA*g* motor to alter its binding to the neighbouring DMT, as proposed for the regulation of IDA*c* by RS2 (ref. [14]). For IDA*f*, we did not observe a direct interaction with IDA*a*, although the TCTEX heterodimer of IDA*f* comes close to the linker of its motor domain, suggesting that they could interact in a beating axoneme.

OIDLs break inter-ODA contacts

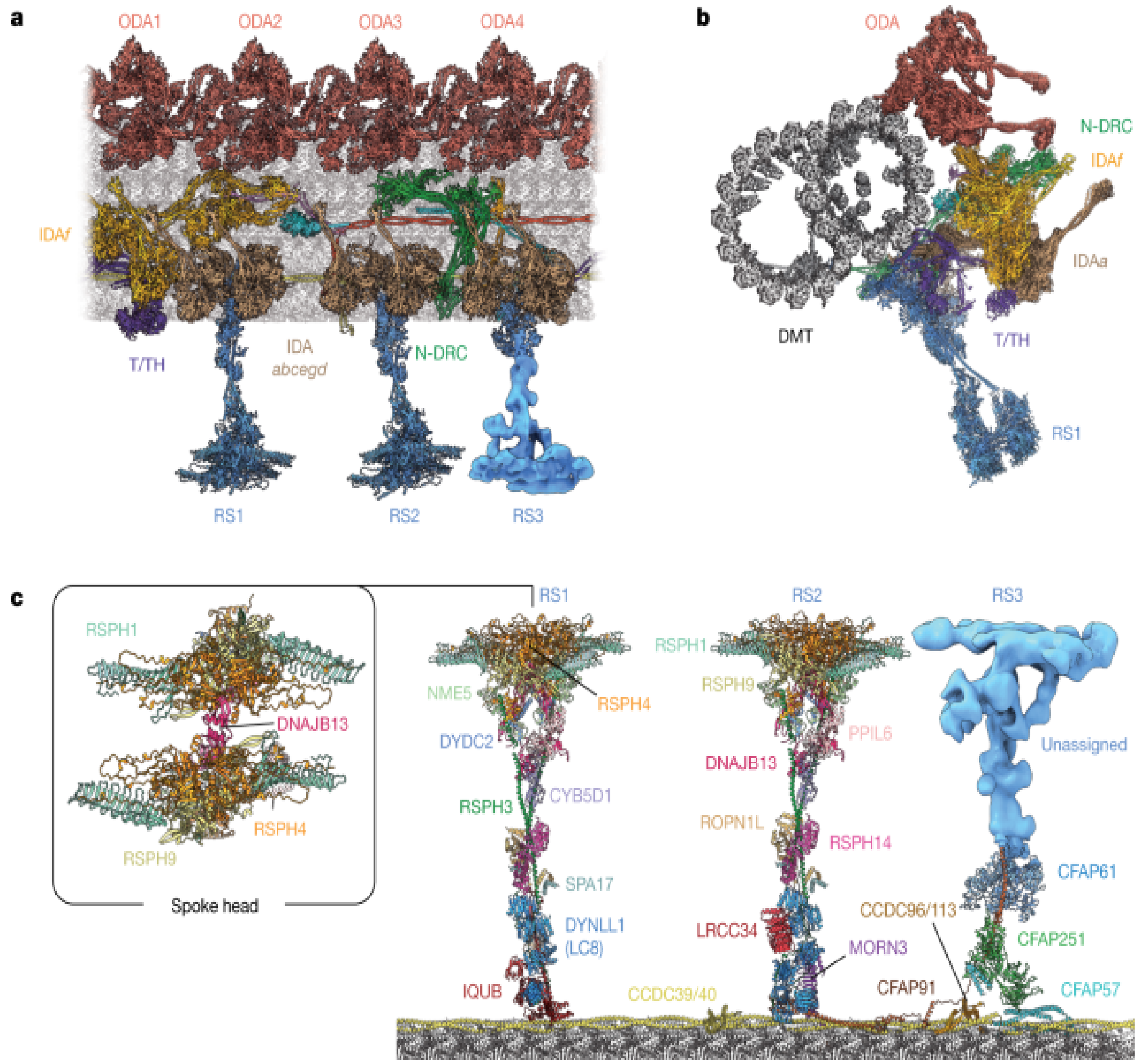
During our analysis of OIDLs, we observed multiple conformations of IDA*f* and N-DRC, consistent with their proposed roles as mechanoregulators of axonemal dynein activity. The core of IDA*f* rotates 6° relative to the DMT (Supplementary Video 2), whereas the N-DRC linker tilts by 9° (Supplementary Video 3). Preservation of the OIDLs during the movement of IDA*f* and N-DRC draws the IC–LC block and γ-HC motor of their partner ODAs forwards by similar amounts (about 20 Å) (Fig. 3f). This movement results in the breaking of inter-ODA contacts between the γ-HC motor and the tail domains of its proximal neighbour into an open state (Supplementary Videos 2 and 3) and also favours separation between the γ- and β-HC tails near the OIDL. We considered that breaking inter-ODA contacts in one complex may send a longitudinal signal that favours the breaking of inter-ODA contacts in adjacent ODA complexes. In support of this hypothesis, we calculated odds ratios between 2.1 and 3.6 for all neighbouring ODA complexes (Fig. 3a, top), indicating that ODA conformation is influenced by the neighbouring ODA. The open and closed states of ODA1 and ODA2 are

highly correlated, with an odds ratio of 3.6 (Fig. 3a), probably through their connections to IDA*f*. However, the conformations of ODA3 and ODA4 with their adjacent ODA complexes are also correlated without a mechanoregulator physically linking them (Fig. 3a). We therefore propose that longitudinal signals caused by the breaking of inter-ODA contacts may activate clusters of ODAs during the propagation of the waveform from the proximal to the distal end of the cilium.

The modular repeat of the human axoneme

To investigate species-specific differences between axonemes, we used our cryo-EM approach to determine a structure and atomic model of the 96-nm modular repeat of human DMT isolated from ciliated 2D organoids resembling the respiratory epithelium (Fig. 4a,b and Supplementary Figs. 13 and 14). The model clarifies the composition of human RS1 and RS2 (Supplementary Tables 3 and 4), revealing simpler spoke heads, necks and bases (in the case of RS1) than in *C. reinhardtii* (Fig. 4c). The simpler architectures of RS1 and RS2 may be compensated by a full-length RS3, which we show is built on a base of CFAP61, CFAP91 and CFAP251 (Fig. 4c and Extended Data Fig. 6a). The homologues of this triumvirate were previously assigned to the calmodulin- and spoke-associated complex[35,36], which we show form the stubby, headless RS3 stand-in (RS3S) in *C. reinhardtii* (Extended Data Fig. 6b,c). Further investigation is required to identify the proteins that bind CFAP61, CFAP91 and CFAP251 to form the rest of RS3 in humans.

Fig. 4: Atomic model of the 96-nm repeat of a human axoneme.



a, Longitudinal section of the 96-nm repeat of a human respiratory-cilia DMT. **b**, Cross-section of a human DMT. **c**, Atomic model of human RSs coloured by subunit. The inset shows the symmetric head of RS1 viewed from above. Only the composition of the base of RS3 is known; the rest of the complex is shown as a density map in **a** and **c**.

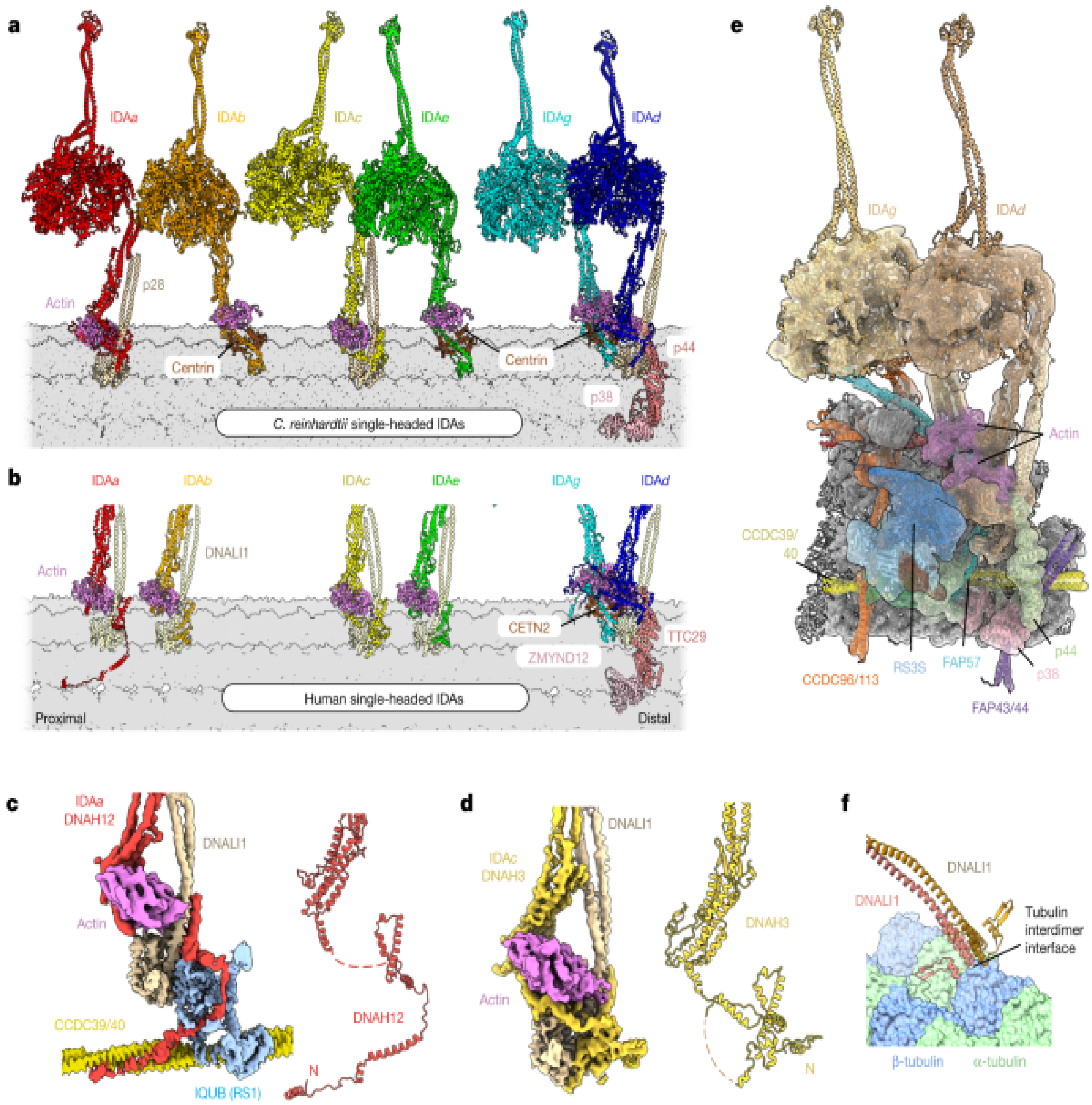
Our model demonstrates that human IDA*f* and the ODA are even more structurally similar than in *C. reinhardtii* because there are fewer additional subunits, including the α-HC and equivalents of LC5 and FAP120. One difference is the direct contact between the TCTEX heterodimer of human IDA*f* and the linker of the IDA*a* motor domain, potentially providing an additional signal-transduction pathway between axonemal dyneins. Unlike ODAs, which have structurally distinct docking complexes in mammals and

algae[5,37], the docking machinery for IDA*f* is highly conserved. The only difference is the absence of a MOT7 equivalent at the centre of the T/TH stellate structure, even though the human genome encodes a potential homologue (TEX47). The site adjacent to ODA4, which is occupied by the distal protrusion in the *C. reinhardtii* axoneme, is vacant (Extended Data Fig. 5e,f), as expected from cryo-ET[32] and the restricted taxonomic distribution of FAP78 homologues (Extended Data Fig. 5g).

Species-dependent IDA organization

An unexpected difference between the axonemes of *C. reinhardtii* and humans is the composition and arrangement of the six single-headed IDAs (Fig. 5a,b). In *C. reinhardtii*, the molecular identity of each HC that forms the IDAs was previously known[38,39], but in humans, the identities of only IDA*f* (DNAH2 and DNAH10), IDA*d* (DNAH1) and IDA*g* (DNAH6) had been confidently deduced on the basis of sequence conservation[40]. By distinguishing structurally distinct HC tails from the cryo-EM density, we have now confirmed those HC assignments[40] and also discovered that IDA*a* is DNAH12, IDA*c* is DNAH3, and IDA*b* and IDA*e* are both DNAH7 (Fig. 5c,d).

Fig. 5: Species-specific compositions and positions of single-headed IDAs.



a, Atomic model of the six single-headed IDAs as they appear on the *C. reinhardtii* DMT. Each IDA is associated with actin and either centrin or a p28 homodimer; p38 and p44 are specific docking factors for IDA*d*. **b**, Cropped image showing the DMT-bound bases of the human single-headed IDAs. Compared with those of *C. reinhardtii*, IDA*b* and IDA*e* have different subunit compositions and docking positions. **c**,**d**, Cryo-EM density of the dynein tail domains was used to identify the HC of IDA*a* as DNAH12 (**c**) and that of IDA*c* as DNAH3 (**d**). **e**, Cryo-EM map and model showing the molecular environment of IDA*g* and IDA*d* and its connection with RS3S in *C. reinhardtii*. A similar arrangement is found at the base of human RS3. **f**,

The DNALI1 homodimer of human IDA*e* interacts directly with the tubulin interdimer interface. Similar interactions with tubulin are observed for human IDA*a,* IDA*b* and IDA*c* and their equivalents in *C. reinhardtii* (through the equivalent p28 homodimer).

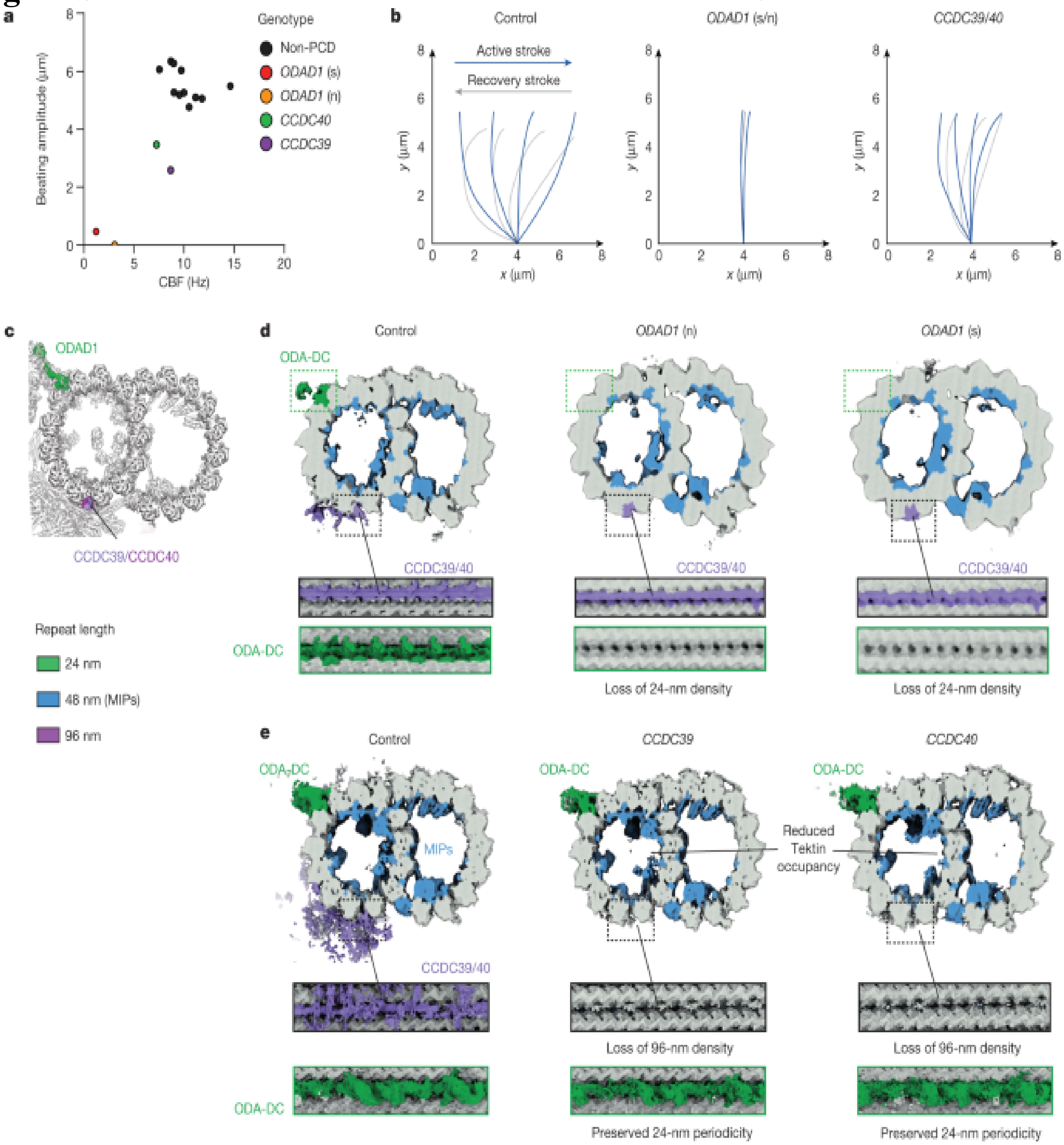
IDA*a*, IDA*c*, IDA*d* and IDA*g* are similar in both organisms. IDA*a* and IDA*c* both have actin and p28/DNALI1 subunits and are directly associated with the microtubule-bound bases of RS1 and RS2, respectively. IDA*d* and IDA*g* are unusual among single-headed IDAs because their N-terminal tails and auxiliary subunits are amalgamated into a single complex that associates with RS3 and ZMYND12/TTC29 complex in humans or the equivalent RS3S and p38/p44 in *C. reinhardtii* (Fig. 5b,e). Thus, the most-conserved IDAs are those that associate with RSs. By contrast, human IDA*b* and IDA*e* differ from their *C. reinhardtii* counterparts in that they have HCs encoded by a single gene, interact with a DNALI1 homodimer instead of centrin and dock onto different sites on the DMT surface. In *C. reinhardtii*, IDA*b* binds at the distal end of IDA*f* on top of the L-shaped molecular ruler (a dimer of the FAP189 family), and IDA*e* binds beside the bulb of the N-DRC over protofilament A04. In humans, IDA*b* and IDA*e* are positioned in line with, and closer to, IDA*a* and IDA*c* (Fig. 5b), where they interact directly with the tubulin surface of protofilament A03 through the DNALI1 subunit (Fig. 5f), rather than with a regulatory complex or docking factor. Despite the different binding location, human IDA*e* still interacts with the N-DRC bulb, which has uncoupled from the baseplate and shifted distally (Extended Data Fig. 4f,g). IDA*e* is lost in *C. reinhardtii* mutants that lack DRC1 or DRC2 (ref. [41]), but whether the N-DRC is required for the docking of human IDA*e* is unknown.

Cryo-EM analysis of mutant axonemes

We then used our cryo-EM approach to investigate how PCD-associated mutations affect human axonemal structure. We prepared axonemes from ciliated organoids derived from four individuals diagnosed with PCD using standard diagnostic criteria (Methods and Extended Data Table 2) and carrying sequence-verified homozygous variants (Extended Data Fig. 7a) selected to disrupt either the 24-nm repeat (by targeting the ODAD1 subunit of the ODA-DC) or the 96-nm repeat (by targeting the CCDC39–CCDC40

(CCDC39/40) coiled coil). All four mutations have been either shown[42] or predicted to create null alleles owing to frameshift or nonsense-mediated degradation, leading to total protein loss. The positions of the affected proteins in the wild-type (control) axoneme are depicted in Fig. 6c.

Fig. 6: Structures of DMTs from individuals with PCD.



a, Beating amplitude is plotted against ciliary beat frequency (CBF) for primary brushing samples from human nasal epithelia for 11 people without PCD and patients with PCD carrying biallelic *ODAD1* (splice site (s) and

nonsense (n)), *CCDC39* and *CCDC40* mutations. **b**, Schematic models of the corresponding ciliary waveforms observed in primary nasal brushings of patients with PCD and people without PCD (control). **c**, Cross-section of the atomic model of the human 96-nm DMT repeat depicting the position of ODAD1 and the CCDC39/40 coiled coil. The proteins are coloured according to their periodicity in the axoneme. **d**,**e**, Cross and longitudinal sections through cryo-EM maps reconstructed from DMTs isolated from pseudostratified respiratory epithelia cultured from either a healthy individual (control) or nasal brushings from patients with PCD. Map densities are coloured according to the repeat length with tubulin shown in grey. **d**, Mutations in *ODAD1* cause complete loss of the ODA-DC (green), which repeats with 24-nm periodicity. **e**, Mutations in *CCDC39* and *CCDC40* cause complete loss of the 96-nm repeat (mauve). Apart from tektins, which had reduced occupancy, most MIPs (blue) are unaffected. Control maps are reconstructed from similar numbers of particles at a similar resolution and periodicity to the mutant maps for direct comparison.

Each mutation caused immotile cilia (more than 90% for *ODAD1* and between 50% and 60% for *CCDC39* and *CCDC40*), with any beating cilia displaying low beat frequency and amplitude (Fig. 6a) and abnormal waveforms (Fig. 6b, Extended Data Fig. 7b and Supplementary Video 4). Clinical transmission electron microscopy (TEM) analysis showed that *ODAD1* mutations caused ODA loss, whereas *CCDC39* and *CCDC40* mutations caused disorganized cross-sections with IDA-deficient DMTs and microtubule transposition (Extended Data Fig. 7c). To better understand how these mutations affect axonemal structure beyond what can be visualized by thin-section TEM, we used cryo-EM to generate structures of their DMTs. For the two *ODAD1* mutants (from patients ID01 and ID02) we could determine structures of only their 8-nm repeats owing to the limited recovery of cilia (Fig. 6d). Comparison with a wild-type DMT structure reconstructed from a similar number of particles to the same resolution and periodicity demonstrated a complete loss of the ODA-DC. Internal MIP structures and the external 96-nm repeat seemed to be unaffected. The structures of the DMTs from the *CCDC39* and *CCDC40* mutants (from patients ID03 and ID04) could be resolved to higher resolution and 48-nm periodicity, the maximum possible, given the complete absence of any recognizable 96-nm repeat, including the CCDC39/40 coiled coil (Fig. 6e). In the absence of regular 96-nm periodicity, RSs are imported into the

axoneme as usual but bind DMTs only sporadically (Extended Data Fig. 7d). In both structures, ODA-DCs and the 24-nm repeat were unaffected but minor defects in the composition of the MIP tektin network were observed.

Collectively, these DMT structures from patients with PCD demonstrate that the loss of docking factors can lead to the complete loss of external periodicities, even when the components (such as RSs) are present in the cilium. The studies also reveal that the 24- and 96-nm external periodicities of the axoneme are independent, as one can be lost without affecting the other, which is consistent with previous observations[43] in *C. reinhardtii*. Furthermore, the loss of external structures had only minor effects on MIP organization, indicating that the establishment of the internal 48-nm repeat is independent of the 96-nm repeat.

Discussion

Our ability to create atomic models of the 96-nm axonemal repeats from two different organisms (*C. reinhardtii* and *Homo sapiens*) demonstrates the strength of combining cryo-EM with AI-based structure prediction for building models of complicated, dynamic cellular machinery. The resulting atomic models reveal that although the axoneme is structurally conserved, there are species-specific differences present in every axonemal complex. The similarities reveal fundamental mechanisms involved in the periodic docking of axonemal complexes onto microtubules and the engagement and manipulation of dyneins by mechanoregulatory sensors. On the other hand, the differences reveal species-specific factors and regulatory mechanisms that fine-tune ciliary waveforms.

Structural diversity of axonemal dyneins

This study comprehensively resolves the structural diversity of axonemal dyneins. We show that multi-headed ODAs and IDA*f* share many structural features, consistent with their proposed evolution by duplication[40]. MIA and the coiled coil of the ODA-DC may have similar evolutionary origins given their apparent functional equivalence as heterodimeric dynein-docking factors. The differences in the subunit composition and binding sites of the single-headed axonemal dyneins highlight the role of these IDAs in the

diversification of axonemal structure and the species-specific fine-tuning of ciliary motility. We show that their divergent N termini are responsible for their localization to specific axonemal substructures, explaining why IDAs are not functionally interchangeable and cannot be substituted by other homologues[3,39]. Our structural information will aid the investigation of dynein HCs in other species because dyneins with structurally conserved N termini are expected to bind to similar positions in the axoneme.

Multipartite docking mechanisms

Our findings provide insights into the complex mechanisms that generate and maintain periodicities across the axoneme. In contrast to previous studies that reinforced the idea that axonemal complexes have single dedicated docking machineries (such as the ODA-DC[5,37]), we demonstrate that multipartite docking mechanisms exist for other axonemal complexes, including IDA*f* (Fig. 2a,b). These docking factors can span large distances and contribute to the binding of multiple axonemal complexes. For example, two of the IDA*f* docking factors, FAP57 and the C-terminal FAP43/44 coiled coil of T/TH, also function in the multipartite docking of IDA*g* and IDA*d* (Fig. 2b, Fig. 5e, and Extended Data Fig. 6d), which explains the partial loss of IDA*g* and IDA*d* in *fap57* mutants[23]. The physical interlocking of shared docking factors seems to provide a robust mechanism to maintain the precise organization and periodicity of the 96-nm repeat over many micrometres. This type of interlocking is also observed for CFAP91, which begins at the base of RS2, interacts with the N-DRC[14] and then links the RS3 subunits CFAP251 and CFAP61 (Fig. 4c).

Insights into axonemal asymmetry

In *C. reinhardtii*, variations in the identity of the coiled-coil docking factors may provide a mechanism for establishing the longitudinal and radial asymmetry of axonemes. We observed that two surface-bound coiled coils, FAP57 and the L-shaped FAP189, have paralogues in *C. reinhardtii* (Extended Data Fig. 8). Our maps probably represent averages of these paralogues in which the quality of the density is determined by sequence conservation among paralogues (Extended Data Fig. 8b–g). We propose that these paralogues are distributed asymmetrically across the axoneme to

regulate IDA composition and function. Consistent with this hypothesis, loss of FAP57 results in partial loss of IDA*g* and IDA*d* from only DMT1 and DMT5–DMT9 (ref. [23]). The functional importance of different paralogues to *C. reinhardtii* motility is seen in the *mbo2* mutant strain that lacks MBO2 (a paralogue of FAP189) and the ability to convert between asymmetrical and symmetrical waveforms[44]. A similar mechanism for establishing asymmetry may operate in humans, who have two orthologues of FAP189 (CFAP58 and CCDC146).

A mechanoregulation mechanism for ODA regulation

By mining our particle dataset for conformational changes, we present a structure-based model for how IDA*f* and N-DRC could regulate ODA activity. In this model, the extensive head-to-tail inter-ODA contacts in the post-powerstroke state constrain dynein conformation, such that this state must be relaxed before ATP turnover and transition to the pre-powerstroke state[6], which has a remodelled set of inter-ODA interactions[45]. Local deformation of the inter-ODA contacts could therefore act to regulate ODA activity by lowering the energy barrier for ATP hydrolysis and facilitating the transition to the pre-powerstroke state, whereas inter-ODA stabilization would be inhibitory. Genetic studies in *C. reinhardtii* favour inhibition. N-DRC and IDA*f* mutations partly suppress the 'paralysed flagella' phenotype caused by lack of RSs or the central apparatus[46,47], and loss of IDA*f* restores the velocity of ODA-driven microtubule disintegration in mutants that lack the central apparatus[48]. However, high-resolution cryo-ET of rapidly frozen, actively beating cilia will be needed to correlate conformational changes in the N-DRC and IDA*f* with dynein activity to determine whether inhibition is relieved by microtubule bending. Our atomic models will help in the interpretation of these cryo-ET studies.

The observed correlation between neighbouring ODA conformations (Fig. 3a) suggests a mechanism for propagating longitudinal signals through the ODA array to create a proximal-to-distal waveform. Crucially, this signal is not dependent on each ODA being directly connected to a mechanoregulator. Connections that are not conserved outside algae, such as the DRC7-mediated OIDL3 (Fig. 3d), may represent a gain of function that

provides a supplemental or redundant mechanism of regulating the ciliary waveform.

As well as regulating ODA behaviour, the N-DRC seems to modulate IDA*g* activity through its conserved interaction with DRC3 in *C. reinhardtii* (Fig. 3e) and humans (Extended Data Fig. 4h). As the N-DRC tilts, the IDA*g* linker and DRC3 form a closer connection (Supplementary Video 3), which could constrain IDA*g* conformation or activity. The functional importance of this interaction can be seen in *C. reinhardtii drc3* mutants, which have an altered flagellar waveform despite having an otherwise fully assembled N-DRC[34], and the formation of hydrocephalus and male infertility in a point mutation of the mouse *drc3* homologue[49]. Overall, our results support a mechanism whereby the N-DRC transmits a mechanical signal through its electrostatic interaction with the neighbouring DMT (Fig. 2e) through a conformational change to ODA3 and IDA*g*.

Cryo-EM as a clinical tool

Our human DMT model contains around 25 proteins linked to PCD and 20 others associated with either PCD-like phenotypes or male infertility (Supplementary Table 5). The model therefore represents a valuable clinical resource for evaluating the likely effect and pathogenicity of disease mutations and identifying new disease candidates for unsolved PCD cohorts. We recommend expanding clinical genetic screening for PCD and male infertility to include all the genes that encode structurally verified proteins of the axoneme. We anticipate that this expansion may help to account for the 20–25% of PCD cases that cannot currently be explained genetically.

Furthermore, our ability to resolve cryo-EM structures of DMTs from individuals provides an experimental approach to examine the impact of PCD-causing mutations on axonemal structure. We have shown that the translational potential of this approach is particularly powerful when combined with molecular genetics. Compared with cross-sectional TEM and even 3D tomography, the higher resolution provided by single-particle cryo-EM allows the detection of minor structural changes, such as perturbation of the tektin network. Cryo-EM could therefore serve as a useful tool to analyse the 30% of PCD cases in which the axonemal ultrastructure seems to

be normal by clinical TEM[50]. Alternatively, by applying our approach to detect variations in cilia across different tissues, we may gain new insights into the molecular basis for the varying susceptibility to dysfunction caused by mutations in specific genes.

Methods

Human subjects

Peripheral blood samples for genetic screening and respiratory epithelial cells, collected from the inferior nasal turbinate by nasal-scrape brushing, were obtained from affected individuals and healthy volunteers under ethical approval granted through the Health Research Authority London Bloomsbury Research Ethics Committee (REC 08/H0713/82; IRAS 103488) and Living Airway Biobank, administered through the UCL Great Ormond Street Institute of Child Health (REC 19/NW/0171, IRAS 261511, Health Research Authority North West Liverpool East Research Ethics Committee). Informed written consent was obtained from all participants before enrolment in the study. Affected individuals were diagnosed with PCD according to European Respiratory Society (ERS) diagnostic guidelines[51].

Human genetic analysis

Genetic screening in patients with PCD was performed as previously described[42] for individual ID01 and by using a targeted next-generation sequencing panel[52] for individuals ID02, ID03 and ID04. A definitive genetic diagnosis was obtained according to ERS diagnostic guidelines[51], involving the identification of biallelic pathogenic variants (found in a homozygous state in each case) in single PCD genes consistent for each defect observed by TEM.

The first individual (ID01) is homozygous for a previously reported c.853G>A substitution in *ODAD1* that causes an intronic insertion that ultimately results in a downstream nonsense codon (p.Ala248Thrfs52*)[42]. The second individual (ID02) carries a homozygous c.448C>T substitution that introduces a premature stop codon in *ODAD1* (p.Arg150*) and has

previously been associated with PCD[53]. The third individual (ID03) has a homozygous exon 1–3 deletion in *CCDC40*, and the fourth individual (ID04) has a homozygous 5-base-pair deletion (c.1964_1968delCTCTT) in *CCDC39* that introduces a premature stop codon through a frameshift (p.Glu655Glyfs*23).

Clinical TEM

TEM of nasal epithelial cells was performed at the University of Leicester, UK. Cells obtained by nasal brushing were fixed in 2.5% glutaraldehyde in either Sorenson's phosphate or 0.05 M sodium cacodylate buffer for more than 24 h. Samples were subsequently incubated in 1% osmium tetroxide before embedding in 2% agar. Samples were dehydrated using an alcohol series that included increasing concentrations of ethanol/methanol and propylene oxide before embedding in epoxy resin. Ultrathin sections were cut and stained with uranyl acetate and Reynold's lead citrate. Images were acquired using a room-temperature JEOL 1400+ transmission electron microscope. Representative examples are provided in Extended Data Fig. 7c.

Cilia beat analysis

Videos of ciliated respiratory epithelium were acquired by high-speed video microscopy and analysed using a program designed in MATLAB R2022a (MathWorks) based on previous publications[54,55]. In brief, positions at the base and tips of cilia at the beginning and end of an active phase of a beat cycle were recorded. Ciliary beat frequency was calculated by fast Fourier transform and video-kymography. This analysis allowed the calculation of ciliary beat pattern parameters, such as cilium length, angle of beating and beating amplitude. Ciliary beat amplitude per second (beating amplitude × ciliary beat frequency) has previously been described as the best approach to discern between individuals with and without PCD[55]. We analysed 104 videos from 11 individuals without PCD as a control.

Epithelial disruption scores

Epithelial disruption scores were calculated from videos of nasal-scrape brushings as described previously[56]. In brief, epithelial edges were identified

in each sample, investigated for protrusions or abnormalities, and scored from 1 to 4, where 1 indicates a normal ciliated edge, 2 indicates a ciliated edge with minor projections, 3 indicates a ciliated edge with major projections, and 4 is for single, unattached, ciliated cells.

Respiratory epithelial culture and cilia isolation

Cells obtained from nasal-scrape brushings (of both a healthy individual and four individuals with PCD) were seeded into collagen-coated wells (PureCol, Sigma-Aldrich, 5006-15MG) containing PneumaCult-Ex Plus medium (STEMCELL Technologies, 05040). The cells were passaged twice: once into a T25 flask and then onto 12 Transwell inserts, each with a surface area of 1.12 cm^2 (Corning, CLS3460) at a density of around 0.5–1.0 million cells per insert. Cells were cultured in PneumaCult-Ex Plus medium until confluent, at which time the medium in the basal chamber was replaced with PneumaCult-ALI medium (STEMCELL Technologies, 05001) and the apical surface was exposed to provide an air–liquid interface (ALI). All media contained Primocin (InvivoGen, ant-pm-05) and penicillin–streptomycin (Thermo Fisher Scientific, 15070063) to prevent the growth of bacteria. Ciliation was observed 4–6 weeks after transition to the ALI. During differentiation, the basolateral medium was refreshed every 2–3 days and the apical surfaces were washed with PBS to remove mucus.

Before deciliation, differentiated ALI cultures were washed twice with PBS for 5 min to remove cell debris and mucus. Ice-cold PBS was then added to both compartments of the culture dish and the dish was placed on ice. After 5 min, the PBS solution was removed and 50 µl ice-cold deciliation buffer (10 mM Tris pH 7.5, 50 mM NaCl, 10 mM $CaCl_2$, 1 mM EDTA, 0.1% Triton X-100, 7 mM β-mercaptoethanol, 1% protease inhibitor cocktail; Sigma-Aldrich, P8340) was added to the cells of each well. After incubation for 2 min without shaking, the cilia-containing solution was transferred to a microcentrifuge tube. This process was repeated six times to ensure that all dissociated cilia were collected. Cellular debris and mucus were removed by centrifugation at 1,000*g* for 1 min at 4 °C. Axonemes were collected by centrifugation at 15,000*g* for 5 min at 4 °C. The axonemal pellet was resuspended in buffer (30 mM HEPES pH 7.3, 1 mM EGTA, 5 mM $MgSO_4$,

0.1 mM EDTA, 25 mM NaCl, 1 mM dithiothreitol, 1% protease inhibitor cocktail), flash-frozen in liquid nitrogen and stored at −80 °C until use.

Sample preparation for cryo-EM

Splayed *C. reinhardtii* axonemes

C. reinhardtii (CC-1690) algae were cultured in standard Tris acetate phosphate medium at room temperature in 12 h:12 h light:dark cycles. Flagella were extracted from *C. reinhardtii* using the dibucaine method[57], as described previously[5]. After removing the cell bodies, the detached flagella in the supernatant were collected by centrifugation at 3,000*g* for 15 min. Freshly isolated flagella were splayed into DMTs as described previously[5]. In brief, flagellar membranes were first removed in fresh HMDEKP (30 mM HEPES, 25 mM KCl, 5 mM $MgSO_4$, 0.5 mM EGTA, Protease Arrest (G-Biosciences)) by adding 1% NP-40 at 4 °C while rotating for 30 min. After washing away the detergent and other soluble proteins, axonemes were splayed at a final concentration of 0.5 mg ml^{-1} by treatment with 10 mM $Mg^{2+}ATP^{2-}$ and 750 µM $CaCl_2$ in HMDEKP while being rotated at room temperature for 1 h. Splayed axonemes were concentrated by centrifugation at 2,500*g* and immediately used to prepare cryo-EM grids. Then 2.5 µl of splayed axoneme sample at 16–26 mg ml^{-1} was applied to glow-discharged C-Flat 1.2/1.3-4Cu grids (Protochips) inside a Vitrobot Mark IV (Thermo Fisher Scientific) kept at 100% humidity. The cryo-EM samples were then incubated 10 s, blotted for 10 s and plunged into liquid ethane cooled by liquid nitrogen.

Human respiratory-cilia axonemes

Human axonemes (3 µl) with an absorbance reading at 280 nm of 5.8 were applied to glow-discharged Quantifoil holey carbon grids (R2/1, copper, 400 mesh, Quantifoil Micro Tools, Q410CR1). The grids were blotted for 10–11 s with a blot force of 10 in 100% humidity before being plunged into liquid ethane using a Vitrobot Mark IV (Thermo Fisher Scientific).

Cryo-EM data collection and processing for splayed *C. reinhardtii* axonemes

A total of 36,918 micrographs of splayed *C. reinhardtii* axonemes were collected using SerialEM as described[58] on a Titan Krios microscope (Thermo Fisher Scientific) under conditions identical to our previous study[5]. Microscope settings are summarized in Extended Data Table 1.

All image processing was done using RELION-3.1 (ref. [59]) or RELION-4.0 (ref. [60]) unless otherwise stated. The dose-fractionated image stacks were aligned and dose-weighted using MotionCor2 software[61]. CTFFIND4 was used to estimate the parameters of the contrast transfer function (CTF)[62]. A total of 31,275 micrographs survived quality control on the basis of the absence of ice contamination, the quality of the Thon rings and the presence of splayed axonemes in the micrographs. Microtubules (including DMTs and singlet microtubules) were manually picked on each micrograph by defining their start and end coordinates in RELION. Overlapping microtubules were avoided. The selected microtubules were computationally divided into overlapping boxes (200 × 200 pixels, bin 2) with an 82-Å non-overlapping region (step size) between adjacent boxes, corresponding to the approximate length of the tubulin α/β-heterodimer. A total of 8,546,747 8-nm particles were extracted and aligned by multiple rounds of 2D classification, with the 106,857 particles that fell into poorly defined classes being discarded. The remaining 8,439,890 particles were analysed using both a DMT-centred approach[13] and an ODA-centred approach[5] (Supplementary Figs. 1–3). In the DMT-centric method, the 8-nm particles underwent 3D refinement and classification to select 2,058,898 particles with well-resolved DMT density. These particles were re-extracted without downscaling in 600-pixel boxes, 3D-refined, CTF-refined, polished and CTF-refined again. In the ODA-centric method, the 8-nm particles underwent 3D refinement and classification to yield three ODA-bound classes. These three classes corresponded to ODA centred in the map and two shifted by ±8 nm from the centre. The off-centre classes were re-extracted to centre the ODA in the map to generate a '24-nm particle' set. To maximize particle number, we extracted new particles 24 nm and 48 nm from this set. A second round of 3D refinement and classification yielded 1,028,815 24-nm particles with a centred ODA. These particles were then re-extracted without downscaling in

600-pixel boxes, 3D-refined, CTF-refined, polished and CTF-refined for a second time.

For both the DMT- and ODA-centric approaches, maps of the 48-nm repeat were obtained by 3D classification using a mask on the lumen of A-tubule protofilaments A08–A13 where MIPs have 48-nm periodicity[13]. For the DMT-centric approach, this yielded six different registries of the 48-nm repeat, and for the ODA-centric approach, two registries from the previously determined 24-nm repeat maps were identified. We then obtained 96-nm maps by a second round of 3D classification using a custom mask over the external axonemal complexes, including RSs, the N-DRC baseplate and IDA*f* to yield a total of 12 different registries from the DMT-centric approach, and 4 different registries from the ODA-centric approach.

To overcome conformational heterogeneity and loss of individual complexes during sample preparation, each axonemal complex was analysed by 3D classification and refinement following the subtraction of signal outside the area of interest (Supplementary Figs. 4–9). Where appropriate, maps were obtained by combining subtracted particles from different registries and from both the DMT- and ODA-centric approaches. When combining particles, strict controls were used to prevent particle duplication.

Focused maps of individual axonemal complexes were positioned on the DMT using unmasked box 600 reference maps from 3D refinement containing IDAs, ODAs and the N-DRC. Because of the low resolution, the motor domains of IDA*b* and IDA*e* are copies of IDA*g*, the other centrin-containing IDA. The maps were then combined into a single composite map that covered a 96-nm section of the DMT using the 'vop maximum' command in ChimeraX[63]. The initial composite map was then cropped using the subregion selection tool of Chimera to remove empty regions and reduce the file size.

Cryo-EM data collection and processing for human respiratory DMTs

A total of 37,071 micrographs of human respiratory DMTs were acquired using a Titan Krios microscope (Thermo Fisher Scientific) under conditions

identical to our previous study[64]. Microscope settings are summarized in Extended Data Table 1. The videos were motion- and CTF-corrected using MotionCor2 (ref. [61]) and CTFFIND4 (ref. [62]), respectively. 16,933 micrographs survived quality control. Microtubule start and end points were selected from the micrographs using RELION manual picking. 2,582,833 8-nm particles were extracted from the micrographs in 512-pixel boxes, downscaled to 256-pixel boxes to accelerate computation. 2D classification was done to check image quality, not to exclude particles.

A 15-Å-resolution map of the bovine respiratory DMT (EMD-24664)[37] was used as an initial reference for 3D refinement, followed by 3D classification without image alignment to exclude bad particles. To increase the number of retained particles, we repeated the process and merged all the good particles while excluding duplicates. In total, 1,903,665 8-nm particles were retained and re-extracted without binning. Then 16-nm particles were generated from the 8-nm particles using 3D classification with a cylindrical mask on 16-nm repeating MIPs near the inner junction. Subsequently 48-nm particles were generated from the 16-nm particles using 3D classification with a cylindrical mask on 48-nm repeating MIPs near the seam of the A tubule. Finally, 96-nm particles were generated from the 48-nm particles using 3D classification with a cylindrical mask on 96-nm repeating axonemal proteins on protofilaments A02–A03. This resulted in two parts of the 96-nm repeat with 90,482 and 87,789 particles. Local refinements using 60 different cylindrical masks, each corresponding to a 16-nm longitudinal section of 2 or 3 protofilaments, were then performed to improve the resolution of the protofilaments and closely associated microtubule-binding proteins. The resolution within each mask improved to 3.3–3.8 Å. The map quality of the external axonemal complexes was improved using particle re-centring, 3D classification, mask-focused local refinement and multi-body refinement. The resulting maps and their nominal resolutions (based on the FSC = 0.143 criterion) are listed in Supplementary Fig. 14. Only 13,340 ODA-containing particles were observed, which prevented analysis of different conformational states, as done for our *C. reinhardtii* dataset. Maps were postprocessed using Phenix_autosharpen and merged in ChimeraX to generate a composite map. A schematic of the processing scheme is provided in Supplementary Fig. 13.

Cryo-EM data of respiratory DMTs from human patients with mutations in *CCDC39* and *CCDC40* were collected on a Titan Krios microscope under conditions identical to those for the wild type, yielding 3,423 and 3,483 micrographs, respectively. Cryo-EM data of respiratory DMTs from human patients with *ODAD1* mutations were collected on a Talos Arctica microscope (Thermo Fisher Scientific) equipped with a K3 detector (Gatan), yielding 56 and 146 micrographs of the *ODAD1* nonsense mutant and the splice mutant, respectively. Data were collected with a defocus range between −1.0 µm and −2.2 µm and at a nominal magnification of ×36,000, yielding a pixel size of 1.1 Å. Each micrograph was fractionated into 50 video frames with a total dose of 52.562 electrons per $Å^2$. Data processing followed the scheme established for the wild-type dataset (Supplementary Fig. 13). For data processing of the *CCDC39* and *CCDC40* mutants, around 280,000 8-nm particles were randomly selected and used to generate 48-nm maps. The same number of wild-type particles were processed in the same way to generate a matched control. The three 48-nm maps were lowpass filtered to 8 Å for comparison in Fig. 6e. For data processing of the *ODAD1* mutants, 12,884 and 35,541 8-nm particles were extracted for the *ODAD1* nonsense mutant and splice mutant, respectively. Only 8-nm maps were generated because of the limited particle numbers. *ODAD1* mutant maps and an 8-nm wild-type map (from the above-mentioned 280,000 particles) were lowpass filtered to 12 Å for comparison in Fig. 6d.

Model building and refinement

Maps of the *C. reinhardtii* DMT were interpreted by combining previous structures, AI-guided structural predictions[65], cryo-ET studies[3,24,39,41,66,67,68] and de novo modelling using Coot[69]. Previous structures were used to model the DMT and its associated MIPs (PDB: 6U42)[13], RS1 and RS2 (PDB: 7JTS, 7JTK and 7JU4)[14], the N-DRC baseplate (PDB: 7JU4)[14] and the ODA with its docking complex (PDB: 7KZM and 7KZO)[5] and LC1 subunit (PDB: 6L4P)[70]. Models of the C1 (PDB: 7SQC)[15] and C2 (PDB: 7SOM)[15] microtubule were used to make Fig. 1b. Structure predictions of other axonemal proteins were obtained using AlphaFold2 (ref. [71]) using primary sequences from Phytozome v.13 and UniProt. Homology models of the dynein motor domains (including the linker, AAA+ modules, stalk and microtubule-binding domain) were generated in the post-powerstroke state

using SWISS-MODEL[72], as they were too large to be modelled by AlphaFold2. Protein–protein interactions were modelled using AlphaFold2 multimer[73]. Examples of subcomplexes modelled using AlphaFold2 multimer are shown in Supplementary Figs. 10–12. Axonemal complexes were assembled from modelled proteins and subcomplexes by positioning them in the cryo-EM maps using ChimeraX[63] or Coot[69]. The subtomogram average of the 96-nm repeat from *C. reinhardtii* (EMD-6872)[24] was used to help position the motor domains of IDA*b* and IDA*e*. Multiple rounds of AlphaFold2 multimer prediction were used to improve the interfaces between neighbouring proteins. Finally, all chains were merged into a single PDB file using Coot and given a unique chain ID. Proteins identified in this study are provided in Supplementary Table 1 and the rationale for their placement is given in Supplementary Table 2.

Our splaying method of sample preparation causes axonemal dyneins to lose their connection with the neighbouring DMT. As a result, we do not observe density for the microtubule-binding domains of the axonemal dyneins, but we included them in our models for completeness. Homology models of motor domains were positioned by fitting into the observed cryo-EM densities for the linker, individual AAA+ modules and, where possible, the base of the stalk. Loss of connection to the neighbouring DMT may have also caused subtle changes in the conformation of the dynein motors relative to intact cilia. Small adjustments of the model may therefore be needed to model subtomogram averages of intact axonemes. In addition, although we observed two different conformational states for the *C. reinhardtii* β-HC, as described previously[5], we have modelled only the orientation in which the stalks extend towards the proximal end, because this conformation is most common in ODA structures from intact axonemes[74].

To construct the model of the axoneme's 9 + 2 architecture (Fig. 1b), we used subtomogram averages of DMT1 (EMD-2113), DMT2-8 (EMD-2132), DMT9 (EMD-2118)[3] and the central apparatus (EMD-31143)[75], and EMD-6756 (ref. [29]) was used to guide the placement of DMTs relative to one another.

The map of the human 96-nm repeat was initially fitted with two copies of the pre-existing model of the human DMT 48-nm repeat (PDB: 7UNG)[64].

This revealed an additional MIP, RIBC1, which had been incorrectly assigned as RIBC2 in the 48-nm repeat. Building of the external axonemal complexes was then guided by the model of the *C. reinhardtii* 96-nm repeat and published models of the bovine ODA-DC (PDB: 7RRO)[37] and mouse RS head (PDB: 7DMP)[76]. Human homologues of *C. reinhardtii* axonemal proteins were identified from the results of BLAST searches (available from http://chlamyfp.org/)[77]. Subcomplexes of the human proteins were predicted by AlphaFold2 multimer and placed in the cryo-EM density. A full list of positioned proteins is provided in Supplementary Table 3. The rationale for the placement of proteins without clear homology is given in Supplementary Table 4.

For both structures, fit-to-map was optimized using real-space refinement implemented in Coot, Namdinator[78] and Phenix[79]. The computational memory requirements needed to refine the 96-nm-long atomic model of the *C. reinhardtii* DMT meant that all-atom refinement could not be performed. Instead, three subregions were refined individually before being merged into a final PDB file. The smaller human model was refined as a single PDB file in Phenix. Atomic displacement factors for both models were refined using Phenix[79]. Side chains of asparagine, glutamine and histidine residues were automatedly flipped to improve hydrogen-bonding networks. Model statistics for the 96-nm modular repeat were generated using phenix.molprobity[80] and are provided in Extended Data Table 1.

Figures

Figure panels and videos displaying cryo-EM maps or atomic models were generated using ChimeraX[63]. Maps coloured by local resolution were generated using RELION. Multiple-sequence alignments and phylogenetic trees were calculated using Clustal Omega[81]. Graphs were plotted in GraphPad Prism (GraphPad Software). Odds ratios (Fig. 3a) were calculated using R v4.0.3. Software used in the project was managed by SBGrid[82].

Materials availability

Genomic samples and sequencing data obtained from patients are stored at University College London. Airway cells are biobanked in the UCL Living

Airway Biobank administered through the UCL Great Ormond Street Institute of Child Health.

Reporting summary

Further information on research design is available in the Nature Portfolio Reporting Summary linked to this article.

Data availability

A composite cryo-EM map of the 96-nm repeat unit of DMTs from *C. reinhardtii* flagella has been deposited to the Electron Microscopy Data Bank (EMDB; https://www.ebi.ac.uk/pdbe/emdb/) with the accession code EMD-40220. The atomic model of the *C. reinhardtii* DMT has been deposited in the Protein Data Bank (PDB; https://www.rcsb.org/) with accession code 8GLV. A composite cryo-EM map of the 96-nm repeat unit of DMTs from human respiratory cilia has been deposited in the EMDB with accession code EMD-35888. The atomic model of the human DMT has been deposited in the PDB with accession code 8J07. For both EMD-40220 and EMD-35888, half maps of the individual axonemal complexes and the masks used for focused refinement have been deposited as additional maps associated with the entries. Source data are provided with this paper.

Code availability

Data and code used to calculate the correlation between open and closed states of the ODA are available at https://doi.org/10.5281/zenodo.6908250.

References

1. Legendre, M., Zaragosi, L.-E. & Mitchison, H. M. Motile cilia and airway disease. *Semin. Cell Dev. Biol.* **110**, 19–33 (2021).

2. Ishikawa, T. Axoneme structure from motile cilia. *Cold Spring Harb. Perspect. Biol.* **9**, a028076 (2017).

3. Bui, K. H., Yagi, T., Yamamoto, R., Kamiya, R. & Ishikawa, T. Polarity and asymmetry in the arrangement of dynein and related structures in the *Chlamydomonas* axoneme. *J. Cell Biol.* **198**, 913–925 (2012).

4. Kubo, S. et al. Remodeling and activation mechanisms of outer arm dyneins revealed by cryo-EM. *EMBO Rep.* **22**, e52911 (2021).

5. Walton, T., Wu, H. & Brown, A. Structure of a microtubule-bound axonemal dynein. *Nat. Commun.* **12**, 477 (2021).

6. Rao, Q. et al. Structures of outer-arm dynein array on microtubule doublet reveal a motor coordination mechanism. *Nat. Struct. Mol. Biol.* **28**, 799–810 (2021).

7. Kagami, O., Takada, S. & Kamiya, R. Microtubule translocation caused by three subspecies of inner-arm dynein from *Chlamydomonas* flagella. *FEBS Lett.* **264**, 179–182 (1990).

8. Sakakibara, H., Kojima, H., Sakai, Y., Katayama, E. & Oiwa, K.Inner-arm dynein c of *Chlamydomonas* flagella is a single-headed processive motor. *Nature* **400**, 586–590 (1999).

9. Kamiya, R. & Yagi, T. Functional diversity of axonemal dyneins as assessed by in vitro and in vivo motility assays of *Chlamydomonas* mutants. *Zool. Sci.* **31**, 633–644 (2014).

10. Lin, J. & Nicastro, D. Asymmetric distribution and spatial switching of dynein activity generates ciliary motility. *Science* **360**, eaar1968 (2018).

11. Viswanadha, R., Sale, W. S. & Porter, M. E. Ciliary motility: regulation of axonemal dynein motors. *Cold Spring Harb. Perspect. Biol.* **9**, a018325 (2017).

12. Kutomi, O. et al. A dynein-associated photoreceptor protein prevents ciliary acclimation to blue light. *Sci. Adv.* **7**, eabf3621 (2021).

13. Ma, M. et al. Structure of the decorated ciliary doublet microtubule. *Cell* **179**, 909–922 (2019).

14. Gui, M. et al. Structures of radial spokes and associated complexes important for ciliary motility. *Nat. Struct. Mol. Biol.* **28**, 29–37 (2021).

15. Gui, M., Wang, X., Dutcher, S. K., Brown, A. & Zhang, R. Ciliary central apparatus structure reveals mechanisms of microtubule patterning. *Nat. Struct. Mol. Biol.* **29**, 483–492 (2022).

16. Smith, E. F. & Sale, W. S. Microtubule binding and translocation by inner dynein arm subtype I1. *Cell Motil. Cytoskeleton* **18**, 258–268 (1991).

17. Kagami, O. & Kamiya, R. Translocation and rotation of microtubules caused by multiple species of *Chlamydomonas* inner-arm dynein. *J. Cell Sci.* **103**, 653–664 (1992).

18. Kotani, N., Sakakibara, H., Burgess, S. A., Kojima, H. & Oiwa, K. Mechanical properties of inner-arm dynein-f (dynein I1) studied with in vitro motility assays. *Biophys. J.* **93**, 886–894 (2007).

19. Wirschell, M. et al. IC97 is a novel intermediate chain of I1 dynein that interacts with tubulin and regulates interdoublet sliding. *Mol. Biol. Cell* **20**, 3044–3054 (2009).

20. Ikeda, K. et al. A novel ankyrin-repeat protein interacts with the regulatory proteins of inner arm dynein f (I1) of *Chlamydomonas reinhardtii*. *Cell Motil. Cytoskeleton* **66**, 448–456 (2009).

21. Yamamoto, R. et al. The MIA complex is a conserved and novel dynein regulator essential for normal ciliary motility. *J. Cell Biol.* **201**, 263–278 (2013).

22. Takada, S. & Kamiya, R. Functional reconstitution of *Chlamydomonas* outer dynein arms from alpha-beta and gamma subunits: requirement of a third factor. *J. Cell Biol.* **126**, 737–745 (1994).

23. Lin, J. et al. FAP57/WDR65 targets assembly of a subset of inner arm dyneins and connects to regulatory hubs in cilia. *Mol. Biol. Cell* **30**, 2659–2680 (2019).

24. Kubo, T., Hou, Y., Cochran, D. A., Witman, G. B. & Oda, T. A microtubule–dynein tethering complex regulates the axonemal inner dynein f (I1). *Mol. Biol. Cell* **29**, 1060–1074 (2018).

25. Fu, G. et al. The I1 dynein-associated tether and tether head complex is a conserved regulator of ciliary motility. *Mol. Biol. Cell* **29**, 1048–1059 (2018).

26. Urbanska, P. et al. Ciliary proteins Fap43 and Fap44 interact with each other and are essential for proper cilia and flagella beating. *Cell. Mol. Life Sci.* **75**, 4479–4493 (2018).

27. Gui, L. et al. Scaffold subunits support associated subunit assembly in the *Chlamydomonas* ciliary nexin–dynein regulatory complex. *Proc. Natl Acad. Sci. USA* **116**, 23152–23162 (2019).

28. Oda, T., Yanagisawa, H. & Kikkawa, M. Detailed structural and biochemical characterization of the nexin-dynein regulatory complex. *Mol. Biol. Cell* **26**, 294–304 (2015).

29. Kubo, T. & Oda, T. Electrostatic interaction between polyglutamylated tubulin and the nexin–dynein regulatory complex regulates flagellar motility. *Mol. Biol. Cell* **28**, 2260–2266 (2017).

30. Heuser, T., Raytchev, M., Krell, J., Porter, M. E. & Nicastro, D. The dynein regulatory complex is the nexin link and a major regulatory node in cilia and flagella. *J. Cell Biol.* **187**, 921–933 (2009).

31. Yamaguchi, H., Oda, T., Kikkawa, M. & Takeda, H. Systematic studies of all PIH proteins in zebrafish reveal their distinct roles in axonemal dynein assembly. *eLife* **7**, e36979 (2018).

32. Lin, J. et al. Cryo-electron tomography reveals ciliary defects underlying human *RSPH1* primary ciliary dyskinesia. *Nat. Commun.* **5**, 5727 (2014).

33. King, S. J. & Dutcher, S. K. Phosphoregulation of an inner dynein arm complex in *Chlamydomonas reinhardtii* is altered in phototactic mutant strains. *J. Cell Biol.* **136**, 177–191 (1997).

34. Awata, J. et al. DRC3 connects the N-DRC to dynein g to regulate flagellar waveform. *Mol. Biol. Cell* **26**, 2788–2800 (2015).

35. Heuser, T., Dymek, E. E., Lin, J., Smith, E. F. & Nicastro, D. The CSC connects three major axonemal complexes involved in dynein regulation. *Mol. Biol. Cell* **23**, 3143–3155 (2012).

36. Dymek, E. E. & Smith, E. F. A conserved CaM- and radial spoke-associated complex mediates regulation of flagellar dynein activity. *J. Cell Biol.* **179**, 515–526 (2007).

37. Gui, M. et al. De novo identification of mammalian ciliary motility proteins using cryo-EM. *Cell* **184**, 5791–5806 (2021).

38. Yagi, T., Uematsu, K., Liu, Z. & Kamiya, R. Identification of dyneins that localize exclusively to the proximal portion of *Chlamydomonas* flagella. *J. Cell Sci.* **122**, 1306–1314 (2009).

39. Bui, K. H., Sakakibara, H., Movassagh, T., Oiwa, K. & Ishikawa, T. Molecular architecture of inner dynein arms in situ in *Chlamydomonas reinhardtii* flagella. *J. Cell Biol.* **183**, 923–932 (2008).

40. Kollmar, M. Fine-tuning motile cilia and flagella: evolution of the dynein motor proteins from plants to humans at high resolution. *Mol. Biol. Evol.* **33**, 3249–3267 (2016).

41. Bower, R. et al. DRC2/CCDC65 is a central hub for assembly of the nexin-dynein regulatory complex and other regulators of ciliary and flagellar motility. *Mol. Biol. Cell* **29**, 137–153 (2018).

42. Onoufriadis, A. et al. Splice-site mutations in the axonemal outer dynein arm docking complex gene *CCDC114* cause primary ciliary dyskinesia. *Am. J. Hum. Genet.* **92**, 88–98 (2013).

43. Oda, T., Yanagisawa, H., Kamiya, R. & Kikkawa, M. A molecular ruler determines the repeat length in eukaryotic cilia and flagella. *Science* **346**, 857–860 (2014).

44. Segal, R. A., Huang, B., Ramanis, Z. & Luck, D. J. Mutant strains of *Chlamydomonas reinhardtii* that move backwards only. *J. Cell Biol.* **98**, 2026–2034 (1984).

45. Movassagh, T., Bui, K. H., Sakakibara, H., Oiwa, K. & Ishikawa, T. Nucleotide-induced global conformational changes of flagellar dynein arms revealed by in situ analysis. *Nat. Struct. Mol. Biol.* **17**, 761–767 (2010).

46. Porter, M. E., Power, J. & Dutcher, S. K. Extragenic suppressors of paralyzed flagellar mutations in *Chlamydomonas reinhardtii* identify loci that alter the inner dynein arms. *J. Cell Biol.* **118**, 1163–1176 (1992).

47. Huang, B., Ramanis, Z. & Luck, D. J. Suppressor mutations in *Chlamydomonas* reveal a regulatory mechanism for flagellar function. *Cell* **28**, 115–124 (1982).

48. Smith, E. F. Regulation of flagellar dynein by the axonemal central apparatus. *Cell Motil. Cytoskeleton* **52**, 33–42 (2002).

49. Ha, S., Lindsay, A. M., Timms, A. E. & Beier, D. R. Mutations in *Dnaaf1* and *Lrrc48* cause hydrocephalus, laterality defects, and sinusitis in mice. *G3* **6**, 2479–2487 (2016).

50. Boon, M. et al. Primary ciliary dyskinesia: critical evaluation of clinical symptoms and diagnosis in patients with normal and abnormal ultrastructure. *Orphanet J. Rare Dis.* **9**, 11 (2014).

51. Lucas, J. S. et al. European Respiratory Society guidelines for the diagnosis of primary ciliary dyskinesia. *Eur. Respir. J.* **49**, 1601090 (2017).

52. Fassad, M. R. et al. Clinical utility of NGS diagnosis and disease stratification in a multiethnic primary ciliary dyskinesia cohort. *J. Med. Genet.* **57**, 322–330 (2020).

53. Davis, S. D. et al. Primary ciliary dyskinesia: longitudinal study of lung disease by ultrastructure defect and genotype. *Am. J. Respir. Crit. Care*

Med. **199**, 190–198 (2019).

54. Blanchon, S. et al. Deep phenotyping, including quantitative ciliary beating parameters, and extensive genotyping in primary ciliary dyskinesia. *J. Med. Genet.* **57**, 237–244 (2020).

55. Papon, J.-F. et al. Quantitative analysis of ciliary beating in primary ciliary dyskinesia: a pilot study. *Orphanet J. Rare Dis.* **7**, 78 (2012).

56. Thomas, B., Rutman, A. & O'Callaghan, C. Disrupted ciliated epithelium shows slower ciliary beat frequency and increased dyskinesia. *Eur. Respir. J.* **34**, 401–404 (2009).

57. Craige, B., Brown, J. M. & Witman, G. B. Isolation of *Chlamydomonas* flagella. *Curr. Protoc. Cell Biol.* **59**, 3.41.1–3.41.9 (2013).

58. Schorb, M., Haberbosch, I., Hagen, W. J. H., Schwab, Y. & Mastronarde, D. N. Software tools for automated transmission electron microscopy. *Nat. Methods* **16**, 471–477 (2019).

59. Zivanov, J. et al. New tools for automated high-resolution cryo-EM structure determination in RELION-3. *eLife* **7**, e42166 (2018).

60. Kimanius, D., Dong, L., Sharov, G., Nakane, T. & Scheres, S. H. W. New tools for automated cryo-EM single-particle analysis in RELION-4.0. *Biochem. J.* **478**, 4169–4185 (2021).

61. Zheng, S. Q. et al. MotionCor2: anisotropic correction of beam-induced motion for improved cryo-electron microscopy. *Nat. Methods* **14**, 331–332 (2017).

62. Rohou, A. & Grigorieff, N. CTFFIND4: fast and accurate defocus estimation from electron micrographs. *J. Struct. Biol.* **192**, 216–221 (2015).

63. Pettersen, E. F. et al. UCSF ChimeraX: structure visualization for researchers, educators, and developers. *Protein Sci.* **30**, 70–82 (2021).

64. Gui, M. et al. SPACA9 is a lumenal protein of human ciliary singlet and doublet microtubules. *Proc. Natl Acad. Sci. USA* **119**, e2207605119 (2022).

65. Mirdita, M. et al. ColabFold: making protein folding accessible to all. *Nat. Methods* **19**, 679–682 (2022).

66. Pigino, G. et al. Cryoelectron tomography of radial spokes in cilia and flagella. *J. Cell Biol.* **195**, 673–687 (2011).

67. Heuser, T. et al. Cryoelectron tomography reveals doublet-specific structures and unique interactions in the I1 dynein. *Proc. Natl Acad. Sci. USA* **109**, E2067–E2076 (2012).

68. Lin, J., Heuser, T., Carbajal-González, B. I., Song, K. & Nicastro, D. The structural heterogeneity of radial spokes in cilia and flagella is conserved. *Cytoskeleton* **69**, 88–100 (2012).

69. Brown, A. et al. Tools for macromolecular model building and refinement into electron cryo-microscopy reconstructions. *Acta Crystallogr. D* **71**, 136–153 (2015).

70. Toda, A., Nishikawa, Y., Tanaka, H., Yagi, T. & Kurisu, G. The complex of outer-arm dynein light chain-1 and the microtubule-binding domain of the γ heavy chain shows how axonemal dynein tunes ciliary beating. *J. Biol. Chem.* **295**, 3982–3989 (2020).

71. Jumper, J. et al. Highly accurate protein structure prediction with AlphaFold. *Nature* **596**, 583–589 (2021).

72. Waterhouse, A. et al. SWISS-MODEL: homology modelling of protein structures and complexes. *Nucleic Acids Res.* **46**, W296–W303 (2018).

73. Evans, R. et al. Protein complex prediction with AlphaFold-Multimer. Preprint at https://doi.org/10.1101/2021.10.04.463034 (2021).

74. Zimmermann, N., Noga, A., Obbineni, J. M. & Ishikawa, T. ATP-induced conformational change of axonemal outer dynein arms studied

by cryo-electron tomography. *EMBO J.* https://doi.org/10.15252/embj.2022112466 (2023).

75. Hou, Y. et al. *Chlamydomonas* FAP70 is a component of the previously uncharacterized ciliary central apparatus projection C2a. *J. Cell Sci.* **134**, jcs258540 (2021).

76. Zheng, W. et al. Distinct architecture and composition of mouse axonemal radial spoke head revealed by cryo-EM. *Proc. Natl Acad. Sci. USA* **118**, e2021180118 (2021).

77. Pazour, G. J., Agrin, N., Leszyk, J. & Witman, G. B. Proteomic analysis of a eukaryotic cilium. *J. Cell Biol.* **170**, 103–113 (2005).

78. Kidmose, R. T. et al. *Namdinator*–automatic molecular dynamics flexible fitting of structural models into cryo-EM and crystallography experimental maps. *IUCrJ* **6**, 526–531 (2019).

79. Afonine, P. V. et al. Real-space refinement in PHENIX for cryo-EM and crystallography. *Acta Crystallogr. D* **74**, 531–544 (2018).

80. Chen, V. B. et al. MolProbity: all-atom structure validation for macromolecular crystallography. *Acta Crystallogr. D* **66**, 12–21 (2010).

81. Sievers, F. et al. Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. *Mol. Syst. Biol.* **7**, 539 (2011).

82. Morin, A. et al. Collaboration gets the most out of software. *eLife* **2**, e01456 (2013).

83. Grotjahn, D. A. et al. Cryo-electron tomography reveals that dynactin recruits a team of dyneins for processive motility. *Nat. Struct. Mol. Biol.* **25**, 203–207 (2018).

84. Bazan, R. et al. Ccdc113/Ccdc96 complex, a novel regulator of ciliary beating that connects radial spoke 3 to dynein g and the nexin link. *PLoS Genet.* **17**, e1009388 (2021).

Acknowledgements

We thank C. Hogg and A. Rutman for TEM of nasal brushings, P. Griffin for acquiring high-speed video microscopy data, and R. K. Rai and P. Griffin for assistance in culturing human airway epithelial cells. Cryo-EM data were collected at the Harvard Center for Cryo-EM with assistance from S. Sterling, R. Walsh and M. Mayer. Cryo-EM data processing was supported by SBGrid (J. O'Connor, S. Rawson and W. Temple) and HMS Research Computing. T.W. was supported by a Helen Hay Whitney postdoctoral fellowship. M.G. was supported by a Charles A. King Trust postdoctoral fellowship. H.M.M. acknowledges funding from NIHR GOSH BRC, the Ministry of Higher Education in Egypt and MRC UCL Confidence in Concept (CiC7). A.B. was supported by grants from the National Institutes of Health (GM141109 and GM143183) and the Pew Charitable Trusts.

Author information

Author notes

1. Miao Gui

   Present address: Liangzhu Laboratory, Zhejiang University, Hangzhou, China

2. These authors contributed equally: Travis Walton, Miao Gui

Authors and Affiliations

1. Department of Biological Chemistry and Molecular Pharmacology, Harvard Medical School, Boston, MA, USA

   Travis Walton, Miao Gui & Alan Brown

2. Genetics and Genomic Medicine Department, UCL Great Ormond Street Institute of Child Health, University College London, London, UK

Simona Velkova, Mahmoud R. Fassad & Hannah M. Mitchison

3. Department of Human Genetics, Medical Research Institute, Alexandria University, Alexandria, Egypt

Mahmoud R. Fassad

4. Centre for PCD Diagnosis and Research, Department of Respiratory Sciences, University of Leicester, Leicester, UK

Robert A. Hirst

5. Department of Pediatric Respiratory Medicine and Allergy, Emma Children's Hospital, Amsterdam University Medical Centers, Amsterdam, The Netherlands

Eric Haarman

6. Infection, Immunity & Inflammation Department, NIHR GOSH BRC, UCL Great Ormond Street Institute of Child Health, University College London, London, UK

Christopher O'Callaghan

7. Royal Brompton Hospital, Guy's and St Thomas' NHS Foundation Trust, London, UK

Mathieu Bottier & Thomas Burgoyne

8. National Heart and Lung Institute, Imperial College London, London, UK

Mathieu Bottier

9. Institute of Ophthalmology, University College London, London, UK

Thomas Burgoyne

Contributions

T.W. performed all work on the *C. reinhardtii* axoneme. M.G. performed all structural work related to human axonemes. T.B. and S.V. prepared human axonemes from cultures. M.R.F. and H.M.M. performed genetic analysis. R.A.H., E.H. and C.O'C. provided clinical phenotype and cilia functional analysis. M.B. performed motility analysis. T.W. and A.B. wrote the paper. All authors approved the final content.

Corresponding author

Correspondence to Alan Brown.

Ethics declarations

Competing interests

The authors declare no competing interests.

Peer review

Peer review information

Nature thanks Haixin Sui and the other, anonymous, reviewer(s) for their contribution to the peer review of this work. Peer reviewer reports are available.

Additional information

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Extended data figures and tables

Extended Data Fig. 1 Structural comparison of inner dynein arm *f* (IDA*f*) and outer dynein arm (ODA) from *C. reinhardtii*.

a, Atomic model showing outer dynein arm 1 (ODA1) and inner dynein arm *f* (IDA*f*) bound to a doublet microtubule (DMT, gray). Dynein subunits are colored based upon their structural or functional similarity and listed in the table. Non-shared subunits are labeled. All other external complexes have been removed for clarity. **b**, Comparison of the cores of ODA (left) and IDA*f* (right). Paralogous subunits in these two structures occupy similar positions. IDA*f*-specific subunits IC97 and FAP120 bind the LC tower, whereas the ODA-specific LC4 binds γ-HC. **c**, Comparison of motor domain orientation in IDA*f* and ODA. The ODA, faded and outlined by a dashed line, is overlaid by IDA*f*. The atomic models were aligned by superposing their cores. α-HC of ODA is omitted for clarity. Compared to ODA, the helical bundles of IDA*f* between the core region and the motors are less curled, resulting in rotation of the motors by 45–65° and displacement by 10 nm. These conformations differ from the active state of cytoplasmic dynein[83], suggesting that the connection between the core and the motor may help specialize each dynein for distinct roles.

Extended Data Fig. 2 The *C. reinhardtii* MIA complex contributes to the docking of inner dynein arm *f* (IDA*f*) to the doublet microtubule (DMT).

a, MIA (a heterodimer of FAP73/FAP100) interacts with protofilament A05 of the DMT through a C-terminal helix of FAP100 beneath the *f*β heavy chain tail. Unmodeled density of FAP100 is also shown. **b**, The N-terminus of the MIA coiled coil interacts with a FAP189 family member near the site of its 90° bend on the DMT surface. **c**, Comparison of the binding of docking complex subunits DC1/2 to ODA and MIA to IDA*f*. Both coiled coils bisect the helical bundles of the dynein heavy chains within the dynein cores. **d**, Cryo-EM map showing the interaction between the *f*α tail, MIA, and the β-propeller domains of a FAP57 homodimer.

Extended Data Fig. 3 Structure of the *C. reinhardtii* tether/tetherhead (T/TH) complex.

a, Composite cryo-EM map colored by T/TH subunit and the motor domains of inner dynein arm (IDA) *f*α and *f*β. The doublet microtubule surface is shown in gray. The N-terminal pair of β-propellers of FAP44 interact with

the AAA1, AAA6, and C-terminal domains of the *f*α motor. **b**, Domain organization of T/TH subunits FAP43, FAP44, and MOT7. Both FAP43 and FAP44 contain an N-terminal pair of β-propellers, followed by a helical stellate structure, and a microtubule-associated coiled coil (CC). MOT7 contains a BLUF domain that is proposed to sense blue light[12]. **c**, Schematic illustrating the six-armed stellate structure of T/TH with MOT7 at the center. **d**, Reverse side of the stellate showing the position of MOT7 and its BLUF domain (bright red) within the cryo-EM map. The resolution is too low to observe any flavin cofactors. On the left, the N-terminal pair of β-propellers from FAP43 bind arm 2 of the stellate. **e**, Interaction of the FAP43/44 coiled coil with the inner junction proteins FAP20 and PACRG. Bulky sidechain residues used for identification are labeled.

Extended Data Fig. 4 Integrative modeling of the nexin-dynein regulatory complex (N-DRC) from *C. reinhardtii* and *H. sapiens*.

a, Domain architecture of *C. reinhardtii* N-DRC subunits. **b**, Architecture of the N-DRC "bulb" in *C. reinhardtii*. **c**, Model of the N-DRC linker in the cryo-ET subtomogram average of the wildtype *C. reinhartii* axoneme (EMD-20338)[27]. Small differences between our model and the cryo-ET map are likely due to the loss of interactions with the neighboring doublet microtubule (DMT) in our splayed axoneme cryo-EM sample. **d**, Model of the N-DRC linker in the subtomogram average of the *drc7* mutant axoneme (EMD-20339)[27]. The DRC7 subunit resides outside of the cryo-ET density, consistent with the loss of DRC7 in the mutant axonemes. Reduced density of the distal lobe, containing the C-termini of DRC9 and 10, suggest an altered or weakened interaction between DRC9/10 and the neighboring DMT in this mutant. **e**, Model of the N-DRC linker in the subtomogram average of the *drc11* mutant axoneme (EMD-20340)[27]. Reduced density of the proximal lobe is consistent with the position of DRC11. **f**, Orthogonal views of the atomic model of human N-DRC. The bulb of human N-DRC (DRC1/9/10) is uncoupled from the baseplate unlike in *C. reinhardtii*. **g**, Interaction between the bulb and the N-terminal tail domain of DNAH7 (IDA*e*). **h**, Interaction between the 3-helix bundle of DRC3 and the linker of DNAH6 (IDA*g*).

Extended Data Fig. 5 Identification of FAP78 as the distal protrusion (DP) of the *C. reinhardtii* axoneme.

a, Domain architecture of FAP78. FAP78 contains 16 armadillo (ARM) repeats interrupted with a helical/unstructured region. The C-terminal half contains a microtubule (MT) binding domain and an EF-hand like domain, which is not resolved in our cryo-EM map. **b**, Cryo-EM map of the DP colored by domain. **c**, Atomic model of FAP78. **d**, Sidechain density of a section of FAP78 boxed in panel c. **e**, The FAP78 model positioned within the subtomogram average of the *C. reinhardtii* axoneme (EMD-6872)[24]. The height of the DP is 155 Å in this map. FAP78 is located within 3–5 nm of IDA*f* and ODA4, suggesting a possible role in dynein regulation, although we do not observe direct interactions. **f**, The DP is absent from axonemes from human respiratory cilia (EMD-5950)[32]. The expected location is denoted by a dashed oval. The subtomogram averages in panels e and f are colored according to the color scheme established in Fig. 1. **g**, Phylogenetic distribution of FAP78 is limited to green algae and other single-celled ciliates. The homolog in *T. thermophila* only aligns to the C-terminal half of FAP78 and does not contain ARM repeats, which likely explains why a distal protrusion is not observed in subtomogram averages of *T. thermophila* axonemes (EMD-12119)[84].

Extended Data Fig. 6 Radial spoke 3 (RS3).

a, Model and cryo-EM map showing the base of human RS3 and its association with IDA*d/g*. Coiled coils from CCDC96/113 and CFAP57 are docking factors for RS3. CFAP43/44 (components of the T/TH complex) are docking factors for IDA*d/g*. **b**, Model and cryo-EM map of *C. reinhardtii* RS3S. **c**, Domain organization of RS3S subunits. The boxes on FAP91 indicate consecutive binding sites for RS2, the N-DRC and RS3S. Both the oxidoreductase-like domain of FAP61 and the calmodulin (CaM)-like domain of FAP251 interact with FAP91. **d**, Atomic model of the docking factors for RS3S in *C. reinhardtii*.

Extended Data Fig. 7 Primary ciliary dyskinesia (PCD) patient information.

a, Integrative Genomics Viewer (IGV) snapshot illustrating the homozygous exon 1–3 deletion in the *CCDC40* gene of patient ID03. **b**, Analysis of high-speed video data from PCD patients showing defects in epithelia consistency, cilia beat frequency, beat angle, and amplitude compared to healthy, non-PCD controls. The box represents the interquartile range, the center represents the median, and the whiskers represent minimum and maximum measurements. Statistical significance was determined by one-way ANOVA (non-parametric Kruskal-Wallis), with * $P \leq 0.05$, ** $P \leq 0.01$, *** $P \leq 0.001$ and **** $P \leq 0.0001$. The number of movies used for the analysis were 104 (control); 3 (*ODAD1* s); 8 (*ODAD1* n); 8 (*CCDC40*), and 6 (*CCDC39*). For each mutation, the movies came from the same patient. *ODAD1* (s) refers to individual ID01 with a splice mutation, and *ODAD1* (n) refers to individual ID02 with a nonsense mutation in *ODAD1*. **c**, Representative clinical transmission electron microscopy (TEM) images selected from a minimum of >100 screened cross-sections showing axoneme cross sections from a non-PCD control (with axonemal complexes labeled), two individuals with mutations in *ODAD1* (ID01 and ID02), an individual with a mutation in *CCDC40* (ID03), and an individual with a mutation in *CCDC39* (ID04). Red arrows mark the absence of outer dynein arms (ODAs). Orange arrows mark microtubule (MT) transposition. Yellow arrows mark the absence of inner dynein arms (IDAs). Purple arrows mark an abnormal central apparatus (CA). **d**, Compared to a non-PCD control, radial spokes (blue arrows) attached to doublet microtubules in individual ID03 (with a *CCDC40* mutation) are intermittently bound, consistent with the loss of strict 96-nm periodicity. These micrographs sections were selected from a total of 37,071 non-PCD control micrographs, and 3,483 *CCDC40* micrographs.

Source data

Extended Data Fig. 8 Coiled-coil docking factors potentially contribute to axonemal asymmetry.

a, Atomic models of the FAP189 and FAP57 homodimers. Other axonemal complexes are omitted for clarity. FAP189 forms an L-shaped coiled coil that occupies the interprotofilament cleft between protofilaments A04 and A05. **b**, Fit of FAP189 atomic model into the cryo-EM map. Labeled amino

acids are denoted below by bold font in the multiple sequence alignment of FAP189 with its paralogs FAP58 and MBO2. The conserved sequence ALSXXLENP, highlighted in red, corresponds to residues that interact with FAP100 of the MIA complex (see Extended Data Fig. 2b). **c**, Agreement of cryo-EM sidechain density with the atomic model of FAP57. Residues labeled in the model are denoted by bold font in the multiple sequence alignment below. **d**, Domain architecture of FAP189, FAP58, and MBO2. **e**, Phylogenetic tree showing the relationship between FAP189, FAP58, and MBO2. FAP189 and FAP58 share 80% sequence identity. **f**, Domain architecture of FAP57, FBB7 and FAP331, which are composed of tandem N-terminal β-propellers and a long, helical C-terminal region. **g**, Phylogenetic tree showing the relationship between FAP57, FBB7, and FAP331.

Extended Data Table 1 Cryo-EM data collection, refinement and validation statistics

Extended Data Table 2 Genetics and diagnostic results for all four PCD cases

Supplementary information

Supplementary Information

Reporting Summary

Peer Review File

Supplementary Video 1

Supplementary Video 2

Supplementary Video 3

Supplementary Video 4

Source data

[Source Data Extended Data Fig. 7]

Rights and permissions

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by/4.0/.

Reprints and Permissions

About this article

Cite this article

Walton, T., Gui, M., Velkova, S. *et al.* Axonemal structures reveal mechanoregulatory and disease mechanisms. *Nature* **618**, 625–633 (2023). https://doi.org/10.1038/s41586-023-06140-2

- Received: 26 July 2022
- Accepted: 27 April 2023
- Published: 31 May 2023
- Issue Date: 15 June 2023
- DOI: https://doi.org/10.1038/s41586-023-06140-2

Share this article

Anyone you share the following link with will be able to read this content:

Get shareable link

Sorry, a shareable link is not currently available for this article.

Copy to clipboard

Provided by the Springer Nature SharedIt content-sharing initiative

A structural model of the core of cilia

Nature Research Briefing 31 May 2023

This article was downloaded by **calibre** from https://www.nature.com/articles/s41586-023-06140-2

Amendments & Corrections

- **Author Correction: The power of genetic diversity in genome-wide association studies of lipids** [26 May 2023]
- **Author Correction: LACC1 bridges NOS2 and polyamine metabolism in inflammatory macrophages** [25 May 2023]
- **Author Correction: Testosterone histories from tusks reveal woolly mammoth musth episodes** [26 May 2023]
- **Author Correction: Short tRNA anticodon stem and mutant eRF1 allow stop codon reassignment** [31 May 2023]
- **Author Correction: Quantifying the cost savings of global solar photovoltaic supply chains** [31 May 2023]

- Author Correction
- Published: 26 May 2023

Author Correction: The power of genetic diversity in genome-wide association studies of lipids

- Sarah E. Graham,
- Shoa L. Clarke,
- Kuan-Han H. Wu,
- Stavroula Kanoni,
- Greg J. M. Zajac,
- Shweta Ramdas,
- Ida Surakka,
- Ioanna Ntalla,
- Sailaja Vedantam,
- Thomas W. Winkler,
- Adam E. Locke,
- Eirini Marouli,
- Mi Yeong Hwang,
- Sohee Han,
- Akira Narita,
- Ananyo Choudhury,
- Amy R. Bentley,
- Kenneth Ekoru,
- Anurag Verma,
- Bhavi Trivedi,
- Hilary C. Martin,
- Karen A. Hunt,
- Qin Hui,
- Derek Klarin,
- Xiang Zhu,

- Gudmar Thorleifsson,
- Anna Helgadottir,
- Daniel F. Gudbjartsson,
- Hilma Holm,
- Isleifur Olafsson,
- Masato Akiyama,
- Saori Sakaue,
- Chikashi Terao,
- Masahiro Kanai,
- Wei Zhou,
- Ben M. Brumpton,
- Humaira Rasheed,
- Sanni E. Ruotsalainen,
- Aki S. Havulinna,
- Yogasudha Veturi,
- QiPing Feng,
- Elisabeth A. Rosenthal,
- Todd Lingren,
- Jennifer Allen Pacheco,
- Sarah A. Pendergrass,
- Jeffrey Haessler,
- Franco Giulianini,
- Yuki Bradford,
- Jason E. Miller,
- Archie Campbell,
- Kuang Lin,
- Iona Y. Millwood,
- George Hindy,
- Asif Rasheed,
- Jessica D. Faul,
- Wei Zhao,
- David R. Weir,
- Constance Turman,
- Hongyan Huang,
- Mariaelisa Graff,
- Anubha Mahajan,
- Michael R. Brown,

- Weihua Zhang,
- Ketian Yu,
- Ellen M. Schmidt,
- Anita Pandit,
- Stefan Gustafsson,
- Xianyong Yin,
- Jian'an Luan,
- Jing-Hua Zhao,
- Fumihiko Matsuda,
- Hye-Mi Jang,
- Kyungheon Yoon,
- Carolina Medina-Gomez,
- Achilleas Pitsillides,
- Jouke Jan Hottenga,
- Gonneke Willemsen,
- Andrew R. Wood,
- Yingji Ji,
- Zishan Gao,
- Simon Haworth,
- Ruth E. Mitchell,
- Jin Fang Chai,
- Mette Aadahl,
- Jie Yao,
- Ani Manichaikul,
- Helen R. Warren,
- Julia Ramirez,
- Jette Bork-Jensen,
- Line L. Kårhus,
- Anuj Goel,
- Maria Sabater-Lleal,
- Raymond Noordam,
- Carlo Sidore,
- Edoardo Fiorillo,
- Aaron F. McDaid,
- Pedro Marques-Vidal,
- Matthias Wielscher,
- Stella Trompet,

- Naveed Sattar,
- Line T. Møllehave,
- Betina H. Thuesen,
- Matthias Munz,
- Lingyao Zeng,
- Jianfeng Huang,
- Bin Yang,
- Alaitz Poveda,
- Azra Kurbasic,
- Claudia Lamina,
- Lukas Forer,
- Markus Scholz,
- Tessel E. Galesloot,
- Jonathan P. Bradfield,
- E. Warwick Daw,
- Joseph M. Zmuda,
- Jonathan S. Mitchell,
- Christian Fuchsberger,
- Henry Christensen,
- Jennifer A. Brody,
- Mary F. Feitosa,
- Mary K. Wojczynski,
- Michael Preuss,
- Massimo Mangino,
- Paraskevi Christofidou,
- Niek Verweij,
- Jan W. Benjamins,
- Jorgen Engmann,
- Rachel L. Kember,
- Roderick C. Slieker,
- Ken Sin Lo,
- Nuno R. Zilhao,
- Phuong Le,
- Marcus E. Kleber,
- Graciela E. Delgado,
- Shaofeng Huo,
- Daisuke D. Ikeda,

- Hiroyuki Iha,
- Jian Yang,
- Jun Liu,
- Hampton L. Leonard,
- Jonathan Marten,
- Börge Schmidt,
- Marina Arendt,
- Laura J. Smyth,
- Marisa Cañadas-Garre,
- Chaolong Wang,
- Masahiro Nakatochi,
- Andrew Wong,
- Nina Hutri-Kähönen,
- Xueling Sim,
- Rui Xia,
- Alicia Huerta-Chagoya,
- Juan Carlos Fernandez-Lopez,
- Valeriya Lyssenko,
- Meraj Ahmed,
- Anne U. Jackson,
- Noha A. Yousri,
- Marguerite R. Irvin,
- Christopher Oldmeadow,
- Han-Na Kim,
- Seungho Ryu,
- Paul R. H. J. Timmers,
- Liubov Arbeeva,
- Rajkumar Dorajoo,
- Leslie A. Lange,
- Xiaoran Chai,
- Gauri Prasad,
- Laura Lorés-Motta,
- Marc Pauper,
- Jirong Long,
- Xiaohui Li,
- Elizabeth Theusch,
- Fumihiko Takeuchi,

- Cassandra N. Spracklen,
- Anu Loukola,
- Sailalitha Bollepalli,
- Sophie C. Warner,
- Ya Xing Wang,
- Wen B. Wei,
- Teresa Nutile,
- Daniela Ruggiero,
- Yun Ju Sung,
- Yi-Jen Hung,
- Shufeng Chen,
- Fangchao Liu,
- Jingyun Yang,
- Katherine A. Kentistou,
- Mathias Gorski,
- Marco Brumat,
- Karina Meidtner,
- Lawrence F. Bielak,
- Jennifer A. Smith,
- Prashantha Hebbar,
- Aliki-Eleni Farmaki,
- Edith Hofer,
- Maoxuan Lin,
- Chao Xue,
- Jifeng Zhang,
- Maria Pina Concas,
- Simona Vaccargiu,
- Peter J. van der Most,
- Niina Pitkänen,
- Brian E. Cade,
- Jiwon Lee,
- Sander W. van der Laan,
- Kumaraswamy Naidu Chitrala,
- Stefan Weiss,
- Martina E. Zimmermann,
- Jong Young Lee,
- Hyeok Sun Choi,

- Maria Nethander,
- Sandra Freitag-Wolf,
- Lorraine Southam,
- Nigel W. Rayner,
- Carol A. Wang,
- Shih-Yi Lin,
- Jun-Sing Wang,
- Christian Couture,
- Leo-Pekka Lyytikäinen,
- Kjell Nikus,
- Gabriel Cuellar-Partida,
- Henrik Vestergaard,
- Bertha Hildalgo,
- Olga Giannakopoulou,
- Qiuyin Cai,
- Morgan O. Obura,
- Jessica van Setten,
- Xiaoyin Li,
- Karen Schwander,
- Natalie Terzikhan,
- Jae Hun Shin,
- Rebecca D. Jackson,
- Alexander P. Reiner,
- Lisa Warsinger Martin,
- Zhengming Chen,
- Liming Li,
- Heather M. Highland,
- Kristin L. Young,
- Takahisa Kawaguchi,
- Joachim Thiery,
- Joshua C. Bis,
- Girish N. Nadkarni,
- Lenore J. Launer,
- Huaixing Li,
- Mike A. Nalls,
- Olli T. Raitakari,
- Sahoko Ichihara,

- Sarah H. Wild,
- Christopher P. Nelson,
- Harry Campbell,
- Susanne Jäger,
- Toru Nabika,
- Fahd Al-Mulla,
- Harri Niinikoski,
- Peter S. Braund,
- Ivana Kolcic,
- Peter Kovacs,
- Tota Giardoglou,
- Tomohiro Katsuya,
- Konain Fatima Bhatti,
- Dominique de Kleijn,
- Gert J. de Borst,
- Eung Kweon Kim,
- Hieab H. H. Adams,
- M. Arfan Ikram,
- Xiaofeng Zhu,
- Folkert W. Asselbergs,
- Adriaan O. Kraaijeveld,
- Joline W. J. Beulens,
- Xiao-Ou Shu,
- Loukianos S. Rallidis,
- Oluf Pedersen,
- Torben Hansen,
- Paul Mitchell,
- Alex W. Hewitt,
- Mika Kähönen,
- Louis Pérusse,
- Claude Bouchard,
- Anke Tönjes,
- Yii-Der Ida Chen,
- Craig E. Pennell,
- Trevor A. Mori,
- Wolfgang Lieb,
- Andre Franke,

- Claes Ohlsson,
- Dan Mellström,
- Yoon Shin Cho,
- Hyejin Lee,
- Jian-Min Yuan,
- Woon-Puay Koh,
- Sang Youl Rhee,
- Jeong-Taek Woo,
- Iris M. Heid,
- Klaus J. Stark,
- Henry Völzke,
- Georg Homuth,
- Michele K. Evans,
- Alan B. Zonderman,
- Ozren Polasek,
- Gerard Pasterkamp,
- Imo E. Hoefer,
- Susan Redline,
- Katja Pahkala,
- Albertine J. Oldehinkel,
- Harold Snieder,
- Ginevra Biino,
- Reinhold Schmidt,
- Helena Schmidt,
- Y. Eugene Chen,
- Stefania Bandinelli,
- George Dedoussis,
- Thangavel Alphonse Thanaraj,
- Sharon L. R. Kardia,
- Norihiro Kato,
- Matthias B. Schulze,
- Giorgia Girotto,
- Bettina Jung,
- Carsten A. Böger,
- Peter K. Joshi,
- David A. Bennett,
- Philip L. De Jager,

- Xiangfeng Lu,
- Vasiliki Mamakou,
- Morris Brown,
- Mark J. Caulfield,
- Patricia B. Munroe,
- Xiuqing Guo,
- Marina Ciullo,
- Jost B. Jonas,
- Nilesh J. Samani,
- Jaakko Kaprio,
- Päivi Pajukanta,
- Linda S. Adair,
- Sonny Augustin Bechayda,
- H. Janaka de Silva,
- Ananda R. Wickremasinghe,
- Ronald M. Krauss,
- Jer-Yuarn Wu,
- Wei Zheng,
- Anneke I. den Hollander,
- Dwaipayan Bharadwaj,
- Adolfo Correa,
- James G. Wilson,
- Lars Lind,
- Chew-Kiat Heng,
- Amanda E. Nelson,
- Yvonne M. Golightly,
- James F. Wilson,
- Brenda Penninx,
- Hyung-Lae Kim,
- John Attia,
- Rodney J. Scott,
- D. C. Rao,
- Donna K. Arnett,
- Steven C. Hunt,
- Mark Walker,
- Heikki A. Koistinen,
- Giriraj R. Chandak,

- Chittaranjan S. Yajnik,
- Josep M. Mercader,
- Teresa Tusié-Luna,
- Carlos A. Aguilar-Salinas,
- Clicerio Gonzalez Villalpando,
- Lorena Orozco,
- Myriam Fornage,
- E. Shyong Tai,
- Rob M. van Dam,
- Terho Lehtimäki,
- Nish Chaturvedi,
- Mitsuhiro Yokota,
- Jianjun Liu,
- Dermot F. Reilly,
- Amy Jayne McKnight,
- Frank Kee,
- Karl-Heinz Jöckel,
- Mark I. McCarthy,
- Colin N. A. Palmer,
- Veronique Vitart,
- Caroline Hayward,
- Eleanor Simonsick,
- Cornelia M. van Duijn,
- Fan Lu,
- Jia Qu,
- Haretsugu Hishigaki,
- Xu Lin,
- Winfried März,
- Esteban J. Parra,
- Miguel Cruz,
- Vilmundur Gudnason,
- Jean-Claude Tardif,
- Guillaume Lettre,
- Leen M. ’t Hart,
- Petra J. M. Elders,
- Scott M. Damrauer,
- Meena Kumari,

- Mika Kivimaki,
- Pim van der Harst,
- Tim D. Spector,
- Ruth J. F. Loos,
- Michael A. Province,
- Bruce M. Psaty,
- Ivan Brandslund,
- Peter P. Pramstaller,
- Kaare Christensen,
- Samuli Ripatti,
- Elisabeth Widén,
- Hakon Hakonarson,
- Struan F. A. Grant,
- Lambertus A. L. M. Kiemeney,
- Jacqueline de Graaf,
- Markus Loeffler,
- Florian Kronenberg,
- Dongfeng Gu,
- Jeanette Erdmann,
- Heribert Schunkert,
- Paul W. Franks,
- Allan Linneberg,
- J. Wouter Jukema,
- Amit V. Khera,
- Minna Männikkö,
- Marjo-Riitta Jarvelin,
- Zoltan Kutalik,
- Francesco Cucca,
- Dennis O. Mook-Kanamori,
- Ko Willems van Dijk,
- Hugh Watkins,
- David P. Strachan,
- Niels Grarup,
- Peter Sever,
- Neil Poulter,
- Jerome I. Rotter,
- Thomas M. Dantoft,

- Fredrik Karpe,
- Matt J. Neville,
- Nicholas J. Timpson,
- Ching-Yu Cheng,
- Tien-Yin Wong,
- Chiea Chuen Khor,
- Charumathi Sabanayagam,
- Annette Peters,
- Christian Gieger,
- Andrew T. Hattersley,
- Nancy L. Pedersen,
- Patrik K. E. Magnusson,
- Dorret I. Boomsma,
- Eco J. C. de Geus,
- L. Adrienne Cupples,
- Joyce B. J. van Meurs,
- Mohsen Ghanbari,
- Penny Gordon-Larsen,
- Wei Huang,
- Young Jin Kim,
- Yasuharu Tabara,
- Nicholas J. Wareham,
- Claudia Langenberg,
- Eleftheria Zeggini,
- Johanna Kuusisto,
- Markku Laakso,
- Erik Ingelsson,
- Goncalo Abecasis,
- John C. Chambers,
- Jaspal S. Kooner,
- Paul S. de Vries,
- Alanna C. Morrison,
- Kari E. North,
- Martha Daviglus,
- Peter Kraft,
- Nicholas G. Martin,
- John B. Whitfield,

- Shahid Abbas,
- Danish Saleheen,
- Robin G. Walters,
- Michael V. Holmes,
- Corri Black,
- Blair H. Smith,
- Anne E. Justice,
- Aris Baras,
- Julie E. Buring,
- Paul M. Ridker,
- Daniel I. Chasman,
- Charles Kooperberg,
- Wei-Qi Wei,
- Gail P. Jarvik,
- Bahram Namjou,
- M. Geoffrey Hayes,
- Marylyn D. Ritchie,
- Pekka Jousilahti,
- Veikko Salomaa,
- Kristian Hveem,
- Bjørn Olav Åsvold,
- Michiaki Kubo,
- Yoichiro Kamatani,
- Yukinori Okada,
- Yoshinori Murakami,
- Unnur Thorsteinsdottir,
- Kari Stefansson,
- Yuk-Lam Ho,
- Julie A. Lynch,
- Daniel J. Rader,
- Philip S. Tsao,
- Kyong-Mi Chang,
- Kelly Cho,
- Christopher J. O’Donnell,
- John M. Gaziano,
- Peter Wilson,
- Charles N. Rotimi,

- Scott Hazelhurst,
- Michèle Ramsay,
- Richard C. Trembath,
- David A. van Heel,
- Gen Tamiya,
- Masayuki Yamamoto,
- Bong-Jo Kim,
- Karen L. Mohlke,
- Timothy M. Frayling,
- Joel N. Hirschhorn,
- Sekar Kathiresan,
- VA Million Veteran Program,
- Global Lipids Genetics Consortium,
- Michael Boehnke,
- Pradeep Natarajan,
- Gina M. Peloso,
- Christopher D. Brown,
- Andrew P. Morris,
- Themistocles L. Assimes ,
- Panos Deloukas,
- Yan V. Sun &
- …
- Cristen J. Willer

Show authors

Nature **volume 618**, pages E19–E20 (2023)

- 1600 Accesses

- 3 Altmetric

- Metrics details

The Original Article was published on 09 December 2021

Correction to: *Nature* https://doi.org/10.1038/s41586-021-04064-3
Published online 9 December 2021

In the version of this article initially published, Noha A. Yousri (Department of Genetic Medicine, Weill Cornell Medicine-Qatar, Doha, Qatar and Department of Computer and Systems Engineering, Alexandria University, Egypt) and Steven C. Hunt (Department of Internal Medicine, University of Utah, Salt Lake City, UT, USA and Department of Genetic Medicine, Weill Cornell Medicine-Qatar, Doha, Qatar) were not included in the author list. In addition, Hieab H. H. Adams (Department of Radiology and Nuclear Medicine, Erasmus MC, University Medical Center Rotterdam, Rotterdam, The Netherlands and Department of Clinical Genetics, Erasmus MC, University Medical Center Rotterdam, Rotterdam, The Netherlands) was shown with an incorrect second affiliation in the HTML and PDF versions of the article. Finally, in the HTML version, Cristen J. Willer was mistakenly listed with an extra affiliation (Princess Al-Jawhara Al-Brahim Centre of Excellence in Research of Hereditary Disorders (PACER-HD), King Abdulaziz University, Jeddah, Saudi Arabia). The authors and affiliations have been corrected in the HTML and PDF versions of the article.

Author information

Author notes

1. These authors contributed equally: Shoa L. Clarke, Kuan-Han H. Wu, Stavroula Kanoni, Greg J. M. Zajac, Shweta Ramdas

2. These authors jointly supervised this work: Gina M. Peloso, Christopher D. Brown, Andrew P. Morris, Themistocles L. Assimes, Panos Deloukas, Yan V. Sun, Cristen J. Willer

Authors and Affiliations

1. Department of Internal Medicine, Division of Cardiology, University of Michigan, Ann Arbor, MI, USA

Sarah E. Graham, Ida Surakka, Chao Xue, Jifeng Zhang, Y. Eugene Chen & Cristen J. Willer

2. VA Palo Alto Health Care System, Palo Alto, CA, USA

Shoa L. Clarke, Xiang Zhu, Philip S. Tsao & Themistocles L. Assimes

3. Department of Medicine, Division of Cardiovascular Medicine, Stanford University School of Medicine, Stanford, CA, USA

Shoa L. Clarke, Erik Ingelsson, Philip S. Tsao & Themistocles L. Assimes

4. Department of Computational Medicine and Bioinformatics, University of Michigan, Ann Arbor, MI, USA

Kuan-Han H. Wu, Wei Zhou & Cristen J. Willer

5. William Harvey Research Institute, Barts and the London School of Medicine and Dentistry, Queen Mary University of London, London, UK

Stavroula Kanoni, Eirini Marouli, Olga Giannakopoulou, Konain Fatima Bhatti & Panos Deloukas

6. Department of Biostatistics and Center for Statistics Genetics, University of Michigan, Ann Arbor, MI, USA

Greg J. M. Zajac, Anne U. Jackson & Michael Boehnke

7. Department of Genetics, Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA, USA

Shweta Ramdas, Anurag Verma & Christopher D. Brown

8. Clinical Pharmacology, William Harvey Research Institute, Barts and The London School of Medicine and Dentistry, Queen Mary University of London, London, UK

Ioanna Ntalla

9. Endocrinology, Boston Children's Hospital, Boston, MA, USA

Sailaja Vedantam & Joel N. Hirschhorn

10. Medical and Population Genetics, Broad Institute, Cambridge, MA, USA

Sailaja Vedantam & Joel N. Hirschhorn

11. Department of Genetic Epidemiology, University of Regensburg, Regensburg, Germany

Thomas W. Winkler, Mathias Gorski, Martina E. Zimmermann, Iris M. Heid & Klaus J. Stark

12. McDonnell Genome Institute and Department of Medicine, Washington University, St Louis, MO, USA

Adam E. Locke

13. Department of Precision Medicine, Division of Genome Science, National Institute of Health, Cheongju-si, South Korea

Mi Yeong Hwang, Sohee Han, Hye-Mi Jang, Kyungheon Yoon, Young Jin Kim & Bong-Jo Kim

14. Tohoku Medical Megabank Organization, Tohoku University, Sendai, Japan

Akira Narita, Gen Tamiya & Masayuki Yamamoto

15. Sydney Brenner Institute for Molecular Bioscience, Faculty of Health Sciences, University of the Witwatersrand, Johannesburg, South Africa

Ananyo Choudhury, Scott Hazelhurst & Michèle Ramsay

16. Center for Research on Genomics and Global Health, National Human Genome Research Institute, National Institutes of Health, Bethesda, MD, USA

Amy R. Bentley, Kenneth Ekoru & Charles N. Rotimi

17. Blizard Institute, Barts and the London School of Medicine and Dentistry, Queen Mary University of London, London, UK

Bhavi Trivedi, Karen A. Hunt & David A. van Heel

18. Wellcome Sanger Institute, Hinxton, UK

Hilary C. Martin & Nigel W. Rayner

19. Department of Epidemiology, Emory University Rollins School of Public Health, Atlanta, GA, USA

Qin Hui & Yan V. Sun

20. Atlanta VA Health Care System, Decatur, GA, USA

Qin Hui & Yan V. Sun

21. Malcolm Randall VA Medical Center, Gainesville, FL, USA

Derek Klarin

22. Division of Vascular Surgery and Endovascular Therapy, University of Florida College of Medicine, Gainesville, FL, USA

Derek Klarin

23. Program in Medical and Population Genetics, Broad Institute of MIT and Harvard, Cambridge, MA, USA

Derek Klarin & Masahiro Kanai

24. Department of Statistics, The Pennsylvania State University, University Park, PA, USA

Xiang Zhu

25. Huck Institutes of the Life Sciences, The Pennsylvania State University, University Park, PA, USA

Xiang Zhu

26. Department of Statistics, Stanford University, Stanford, CA, USA

Xiang Zhu

27. deCODE genetics/Amgen, Reykjavik, Iceland

Gudmar Thorleifsson, Anna Helgadottir, Daniel F. Gudbjartsson, Hilma Holm, Unnur Thorsteinsdottir & Kari Stefansson

28. School of Engineering and Natural Sciences, University of Iceland, Reykjavik, Iceland

Daniel F. Gudbjartsson

29. Department of Clinical Biochemistry, Landspitali–National University Hospital of Iceland, Reykjavik, Iceland

Isleifur Olafsson

30. Laboratory for Statistical Analysis, RIKEN Center for Integrative Medical Sciences, Yokohama, Japan

Masato Akiyama, Saori Sakaue, Masahiro Kanai, Yoichiro Kamatani & Yukinori Okada

31. Department of Ophthalmology, Graduate School of Medical Sciences, Kyushu University, Fukuoka, Japan

Masato Akiyama

32. Department of Statistical Genetics, Osaka University Graduate School of Medicine, Osaka, Japan

Saori Sakaue & Yukinori Okada

33. Department of Allergy and Rheumatology, Graduate School of Medicine, The University of Tokyo, Tokyo, Japan

Saori Sakaue

34. Laboratory for Statistical and Translational Genetics, RIKEN Center for Integrative Medical Sciences, Yokohama, Japan

Chikashi Terao

35. Department of Biomedical Informatics, Harvard Medical School, Boston, MA, USA

Masahiro Kanai

36. Analytic and Translational Genetics Unit, Massachusetts General Hospital, Boston, MA, USA

Wei Zhou

37. Program in Medical and Population Genetics, Broad Institute of Harvard and MIT, Cambridge, MA, USA

Wei Zhou & Pradeep Natarajan

38. K.G. Jebsen Center for Genetic Epidemiology, Department of Public Health and Nursing, NTNU, Norwegian University of Science and Technology, Trondheim, Norway

Ben M. Brumpton, Humaira Rasheed, Kristian Hveem & Bjørn Olav Åsvold

39. MRC Integrative Epidemiology Unit, University of Bristol, Bristol, UK

Ben M. Brumpton & Humaira Rasheed

40. Department of Thoracic Medicine, St Olavs Hospital, Trondheim University Hospital, Trondheim, Norway

Ben M. Brumpton

41. Institute for Molecular Medicine Finland (FIMM), HiLIFE, University of Helsinki, Helsinki, Finland

Sanni E. Ruotsalainen, Aki S. Havulinna, Anu Loukola, Sailalitha Bollepalli, Jaakko Kaprio, Samuli Ripatti & Elisabeth Widén

42. Department of Public Health and Welfare, Finnish Institute for Health and Welfare, Helsinki, Finland

Aki S. Havulinna, Pekka Jousilahti & Veikko Salomaa

43. Department of Genetics, Institute for Biomedical Informatics, Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA, USA

Yogasudha Veturi, Yuki Bradford, Jason E. Miller & Marylyn D. Ritchie

44. Division of Clinical Pharmacology, Department of Medicine, Vanderbilt University Medical Center, Nashville, TN, USA

QiPing Feng

45. Department of Medicine (Medical Genetics), University of Washington, Seattle, WA, USA

Elisabeth A. Rosenthal

46. Division of Biomedical Informatics, Cincinnati Children's Hospital Medical Center, Cincinnati, OH, USA

Todd Lingren

47. Center for Genetic Medicine, Northwestern University, Evanston, IL, USA

Jennifer Allen Pacheco

48. Genentech, South San Francisco, CA, USA

Sarah A. Pendergrass & Mark I. McCarthy

49. Fred Hutchinson Cancer Research Center, Division of Public Health Sciences, Seattle, WA, USA

Jeffrey Haessler & Charles Kooperberg

50. Division of Preventive Medicine, Brigham and Women's Hospital, Boston, MA, USA

Franco Giulianini, Julie E. Buring, Paul M. Ridker & Daniel I. Chasman

51. Centre for Genomic and Experimental Medicine, Institute of Genetics and Molecular Medicine, University of Edinburgh, Western General Hospital, Edinburgh, UK

Archie Campbell

52. Usher Institute for Population Health Sciences and Informatics, The University of Edinburgh, Edinburgh, UK

Archie Campbell

53. Clinical Trial Service Unit and Epidemiological Studies Unit, Nuffield Department of Population Health, University of Oxford, Oxford, UK

Kuang Lin, Iona Y. Millwood, Zhengming Chen, Robin G. Walters & Michael V. Holmes

54. Medical Research Council Population Health Research Unit, Nuffield Department of Population Health, University of Oxford, Oxford, UK

Iona Y. Millwood, Zhengming Chen, Robin G. Walters & Michael V. Holmes

55. Department of Population Medicine, Qatar University College of Medicine, QU Health, Doha, Qatar

George Hindy

56. Center for Non-Communicable Diseases, Karachi, Pakistan

Asif Rasheed, Shahid Abbas & Danish Saleheen

57. Survey Research Center, Institute for Social Research, University of Michigan, Ann Arbor, MI, USA

Jessica D. Faul, David R. Weir & Jennifer A. Smith

58. Department of Epidemiology, School of Public Health, University of Michigan, Ann Arbor, MI, USA

Wei Zhao

59. Program in Genetic Epidemiology and Statistical Genetics, Department of Epidemiology, Harvard T. H. Chan School of Public Health, Boston, MA, USA

Constance Turman, Hongyan Huang & Peter Kraft

60. Department of Epidemiology, University of North Carolina, Chapel Hill, NC, USA

Mariaelisa Graff, Heather M. Highland, Kristin L. Young & Kari E. North

61. Wellcome Centre for Human Genetics, University of Oxford, Oxford, UK

Anubha Mahajan, Anuj Goel, Nigel W. Rayner, Mark I. McCarthy & Hugh Watkins

62. Human Genetics Center, Department of Epidemiology, Human Genetics, and Environmental Sciences, School of Public Health, The University of Texas Health Science Center at Houston, Houston, TX, USA

Michael R. Brown, Paul S. de Vries & Alanna C. Morrison

63. Department of Epidemiology and Biostatistics, Imperial College London, London, UK

Weihua Zhang & John C. Chambers

64. Department of Cardiology, Ealing Hospital, London North West University Healthcare NHS Trust, Southall, UK

Weihua Zhang, John C. Chambers & Jaspal S. Kooner

65. Imperial College Healthcare NHS Trust, London, UK

Weihua Zhang & Jaspal S. Kooner

66. Department of Biostatistics, Center for Statistical Genetics, University of Michigan, Ann Arbor, MI, USA

Ketian Yu, Ellen M. Schmidt, Anita Pandit, Xianyong Yin & Goncalo Abecasis

67. Department of Medical Sciences, Molecular Epidemiology and Science for Life Laboratory, Uppsala University, Uppsala, Sweden

Stefan Gustafsson & Erik Ingelsson

68. MRC Epidemiology Unit, University of Cambridge School of Clinical Medicine, Cambridge, UK

Jian’an Luan, Nicholas J. Wareham & Claudia Langenberg

69. Cardiovascular Epidemiology Unit, Department of Public Health and Primary Care, University of Cambridge Strangeways Research

Laboratory, Cambridge, UK

Jing-Hua Zhao

70. Center for Genomic Medicine, Kyoto University Graduate School of Medicine, Kyoto, Japan

Fumihiko Matsuda, Takahisa Kawaguchi & Yasuharu Tabara

71. Department of Internal Medicine, Erasmus MC, University Medical Center Rotterdam, Rotterdam, The Netherlands

Carolina Medina-Gomez & Joyce B. J. van Meurs

72. Department of Epidemiology, Erasmus MC, University Medical Center Rotterdam, Rotterdam, The Netherlands

Carolina Medina-Gomez, Natalie Terzikhan, M. Arfan Ikram, Joyce B. J. van Meurs & Mohsen Ghanbari

73. Department of Biostatistics, Boston University School of Public Health, Boston, MA, USA

Achilleas Pitsillides, L. Adrienne Cupples & Gina M. Peloso

74. Department of Biological Psychology, Behavioral and Movement Sciences, Vrije Universiteit Amsterdam, Amsterdam, The Netherlands

Jouke Jan Hottenga, Gonneke Willemsen, Dorret I. Boomsma & Eco J. C. de Geus

75. Amsterdam Public Health, VU Medical Center Amsterdam, Amsterdam, The Netherlands

Jouke Jan Hottenga

76. Amsterdam Public Health Research Institute, VU Medical Center Amsterdam, Amsterdam, The Netherlands

Gonneke Willemsen

77. Genetics of Complex Traits, University of Exeter Medical School, University of Exeter, Exeter, UK

Andrew R. Wood, Yingji Ji & Timothy M. Frayling

78. Department of Clinical Acupuncture and Moxibustion, Nanjing University of Chinese Medicine, Nanjing, China

Zishan Gao

79. Research Unit of Molecular Epidemiology, Helmholtz Zentrum München, German Research Center for Environmental Health, Neuherberg, Germany

Zishan Gao & Christian Gieger

80. Institute of Epidemiology, Helmholtz Zentrum München, German Research Center for Environmental Health, Neuherberg, Germany

Zishan Gao, Annette Peters & Christian Gieger

81. MRC Integrative Epidemiology Unit at the University of Bristol, Bristol, UK

Simon Haworth, Ruth E. Mitchell & Nicholas J. Timpson

82. Bristol Dental School, University of Bristol, Bristol, UK

Simon Haworth

83. Population Health Sciences, Bristol Medical School, University of Bristol, Bristol, UK

Ruth E. Mitchell & Nicholas J. Timpson

84. Saw Swee Hock School of Public Health, National University of Singapore and National University Health System, Singapore,

Singapore

Jin Fang Chai, Xueling Sim, E. Shyong Tai & Rob M. van Dam

85. Center for Clinical Research and Prevention, Bispebjerg and Frederiksberg Hospital, Copenhagen, Denmark

Mette Aadahl, Line L. Kårhus, Line T. Møllehave, Betina H. Thuesen, Allan Linneberg & Thomas M. Dantoft

86. The Institute for Translational Genomics and Population Sciences, Department of Pediatrics, Lundquist Institute for Biomedical Innovations (Formerly LABioMed) at Harbor-UCLA Medical Center, Torrance, CA, USA

Jie Yao, Xiaohui Li, Yii-Der Ida Chen, Xiuqing Guo & Jerome I. Rotter

87. Center for Public Health Genomics, University of Virginia, Charlottesville, VA, USA

Ani Manichaikul

88. William Harvey Research Institute, Barts and The London School of Medicine and Dentistry, John Vane Science Centre, Queen Mary University of London, London, UK

Helen R. Warren, Julia Ramirez, Mark J. Caulfield & Patricia B. Munroe

89. NIHR Barts Cardiovascular Biomedical Research Centre, Barts and The London School of Medicine and Dentistry, Queen Mary University of London, London, UK

Helen R. Warren, Morris Brown, Mark J. Caulfield, Patricia B. Munroe & Panos Deloukas

90. Novo Nordisk Foundation Center for Basic Metabolic Research, Faculty of Health and Medical Sciences, University of Copenhagen,

Copenhagen, Denmark

Jette Bork-Jensen, Henrik Vestergaard, Oluf Pedersen, Torben Hansen & Niels Grarup

91. Division of Cardiovascular Medicine, Radcliffe Department of Medicine, John Radcliffe Hospital, University of Oxford, Oxford, UK

Anuj Goel

92. Group of Genomics of Complex Diseases, Research Institute of Hospital de la Santa Creu i Sant Pau (IIB Sant Pau), Barcelona, Spain

Maria Sabater-Lleal

93. Cardiovascular Medicine Unit, Department of Medicine, Karolinska Institutet, Center for Molecular Medicine, Karolinska University Hospital, Stockholm, Sweden

Maria Sabater-Lleal

94. Department of Internal Medicine, Section Gerontology and Geriatrics, Leiden University Medical Center, Leiden, The Netherlands

Raymond Noordam

95. Institute for Genetic and Biomedical Research, Italian National Council of Research (IRGB CNR), Cagliari,, Italy

Carlo Sidore

96. Institute for Genetic and Biomedical Research, Italian National Council of Research (IRGB CNR), Lanusei, Italy

Edoardo Fiorillo

97. University Center for Primary Care and Public Health, University of Lausanne, Lausanne, Switzerland

Aaron F. McDaid

98. Swiss Institute of Bioinformatics, Lausanne, Switzerland

Aaron F. McDaid & Zoltan Kutalik

99. Department of Medicine, Internal Medicine, Lausanne University Hospital and University of Lausanne, Lausanne, Switzerland

Pedro Marques-Vidal

100. Department of Epidemiology and Biostatistics, MRC–PHE Centre for Environment and Health, School of Public Health, Imperial College London, London, UK

Matthias Wielscher & Marjo-Riitta Jarvelin

101. Department of Cardiology, Leiden University Medical Center, Leiden, The Netherlands

Stella Trompet & J. Wouter Jukema

102. Department of Internal Medicine, Section of Gerontology and Geriatrics, Leiden University Medical Center, Leiden, The Netherlands

Stella Trompet

103. BHF Glasgow Cardiovascular Research Centre, Faculty of Medicine, Glasgow, UK

Naveed Sattar

104. Institute for Cardiogenetics, University of Lübeck, DZHK (German Research Centre for Cardiovascular Research), Partner site Hamburg/Lübeck/Kiel, University Heart Center Lübeck, Lübeck, Germany

Matthias Munz

105. Charité–University Medicine Berlin, Member of Freie Universität Berlin and Humboldt-Universität zu Berlin, Berlin, Germany

Matthias Munz

106. Department of Periodontology and Synoptic Dentistry, Berlin Institute of Health, Institute for Dental and Craniofacial Sciences, Berlin, Germany

Matthias Munz

107. Deutsches Herzzentrum München, Klinik für Herz- und Kreislauferkrankungen, Technische Universität München, Munich, Germany

Lingyao Zeng

108. Deutsches Zentrum für Herz-Kreislauf-Forschung (DZHK), Partner site Munich Heart Alliance, Munich, Germany

Lingyao Zeng

109. Key Laboratory of Cardiovascular Epidemiology and Department of Epidemiology, State Key Laboratory of Cardiovascular Disease, Fuwai Hospital, National Center for Cardiovascular Diseases, Chinese Academy of Medical Sciences and Peking Union Medical College, Beijing, China

Jianfeng Huang, Bin Yang, Shufeng Chen, Fangchao Liu, Xiangfeng Lu & Dongfeng Gu

110. Lund University Diabetes Centre, Malmo, Sweden

Alaitz Poveda, Azra Kurbasic & Paul W. Franks

111. Department of Genetics and Pharmacology, Institute of Genetic Epidemiology, Medical University of Innsbruck, Innsbruck, Austria

Claudia Lamina & Lukas Forer

112. German Chronic Kidney Disease Study, Innsbruck, Austria

Claudia Lamina, Lukas Forer & Florian Kronenberg

113. Institute for Medical Informatics, Statistics and Epidemiology, University of Leipzig, Leipzig, Germany

Markus Scholz & Markus Loeffler

114. LIFE Research Centre for Civilization Diseases, University of Leipzig, Leipzig, Germany

Markus Scholz, Joachim Thiery & Markus Loeffler

115. Radboud University Medical Center, Radboud Institute for Health Sciences, Nijmegen, The Netherlands

Tessel E. Galesloot, Lambertus A. L. M. Kiemeney & Jacqueline de Graaf

116. Quantinuum Research, Wayne, PA, USA

Jonathan P. Bradfield

117. Division of Statistical Genomics, Department of Genetics, Washington University School of Medicine, St Louis, MO, USA

E. Warwick Daw, Mary F. Feitosa, Mary K. Wojczynski, Karen Schwander & Michael A. Province

118. Department of Epidemiology, University of Pittsburgh, Pittsburgh, PA, USA

Joseph M. Zmuda

119. Institute for Biomedicine, Eurac Research, Affiliated Institute of the University of Lübeck, Bolzano, Italy

Jonathan S. Mitchell, Christian Fuchsberger & Peter P. Pramstaller

120. Department of Clinical Biochemistry, Lillebaelt Hospital, Vejle, Denmark

Henry Christensen & Ivan Brandslund

121. Cardiovascular Health Research Unit, Department of Medicine, University of Washington, Seattle, WA, USA

Jennifer A. Brody & Joshua C. Bis

122. The Charles Bronfman Institute for Personalized Medicine, Icahn School of Medicine at Mount Sinai, New York, NY, USA

Michael Preuss, Girish N. Nadkarni & Ruth J. F. Loos

123. Department of Twin Research and Genetic Epidemiology, King’s College London, London, UK

Massimo Mangino, Paraskevi Christofidou & Tim D. Spector

124. NIHR Biomedical Research Centre at Guy’s and St Thomas’ Foundation Trust, London, UK

Massimo Mangino

125. Department of Cardiology, University Medical Center Groningen, University of Groningen, Groningen, The Netherlands

Niek Verweij, Jan W. Benjamins & Pim van der Harst

126. Institute of Cardiovascular Sciences, University College London, London, UK

Jorgen Engmann

127. Department of Epidemiology and Public Health, University College London, London, UK

Jorgen Engmann & Mika Kivimaki

128. Department of Psychiatry, University of Pennsylvania, Philadelphia, PA, USA

Rachel L. Kember

129. Amsterdam UMC, Department of Epidemiology and Biostatistics, Amsterdam Public Health Research Institute, Amsterdam, The Netherlands

Roderick C. Slieker, Morgan O. Obura, Joline W. J. Beulens & Leen M. ’t Hart

130. Department of Cell and Chemical Biology, Leiden University Medical Center, Leiden, The Netherlands

Roderick C. Slieker & Leen M. ’t Hart

131. Montreal Heart Institute, Montreal, Quebec, Canada

Ken Sin Lo, Jean-Claude Tardif & Guillaume Lettre

132. Icelandic Heart Association, Reykjavik, Iceland

Nuno R. Zilhao & Vilmundur Gudnason

133. Department of Anthropology, University of Toronto at Mississauga, Mississauga, Ontario, Canada

Phuong Le & Esteban J. Parra

134. Vth Department of Medicine, Medical Faculty Mannheim, Heidelberg University, Mannheim, Germany

Marcus E. Kleber, Graciela E. Delgado & Winfried März

135. Synlab MVZ Humangenetik Mannheim, Mannheim, Germany

Marcus E. Kleber

136. Biomedical Technology Research Center, Tokushima Research Institute, Otsuka Pharmaceutical, Tokushima, Japan

Daisuke D. Ikeda, Hiroyuki Iha & Haretsugu Hishigaki

137. Institute for Molecular Bioscience, The University of Queensland, Brisbane, Queensland, Australia

Jian Yang

138. Institute for Advanced Research, Wenzhou Medical University, Wenzhou, China

Jian Yang

139. Nuffield Department of Population Health, University of Oxford, Oxford, UK

Jun Liu & Cornelia M. van Duijn

140. Laboratory of Neurogenetics, National Institute on Aging, NIH, Bethesda, MD, USA

Hampton L. Leonard & Mike A. Nalls

141. Data Tecnica International, Glen Echo, MD, USA

Hampton L. Leonard & Mike A. Nalls

142. MRC Human Genetics Unit, Institute of Genetics and Cancer, University of Edinburgh, Western General Hospital, Edinburgh, UK

Jonathan Marten, Paul R. H. J. Timmers, James F. Wilson, Veronique Vitart & Caroline Hayward

143. Institute for Medical Informatics, Biometrie and Epidemiology, University of Duisburg-Essen, Essen, Germany

Börge Schmidt, Marina Arendt & Karl-Heinz Jöckel

144. Department of Computer Science, University of Applied Sciences and Arts Dortmund, Dortmund, Germany

Marina Arendt

145. Centre for Public Health, Queen's University of Belfast, Belfast, UK

Laura J. Smyth, Marisa Cañadas-Garre, Amy Jayne McKnight & Frank Kee

146. Department of Epidemiology and Biostatistics, School of Public Health, Tongji Medical College, Huazhong University of Science and Technology, Wuhan, China

Chaolong Wang

147. Genome Institute of Singapore, Agency for Science, Technology and Research, Singapore, Singapore

Chaolong Wang, Rajkumar Dorajoo, Jianjun Liu & Chiea Chuen Khor

148. Public Health Informatics Unit, Department of Integrated Health Sciences, Nagoya University Graduate School of Medicine, Nagoya, Japan

Masahiro Nakatochi

149. MRC Unit for Lifelong Health and Ageing at UCL, London, UK

Andrew Wong & Nish Chaturvedi

150. Department of Pediatrics, Tampere University Hospital, Tampere, Finland

Nina Hutri-Kähönen

151. Department of Pediatrics, Faculty of Medicine and Health Technology, Tampere University, Tampere, Finland

Nina Hutri-Kähönen

152. Brown Foundation Institute of Molecular Medicine, McGovern Medical School, University of Texas Health Science Center at Houston, Houston, TX, USA

Rui Xia & Myriam Fornage

153. Departamento de Medicina Genómica y Toxicología Ambiental, Instituto de Investigaciones Biomédicas, UNAM, Ciudad de Mexico, Mexico, Mexico

Alicia Huerta-Chagoya

154. Departamento de Genómica Computacional, Instituto Nacional de Medicina Genómica, Ciudad de Mexico, Mexico, Mexico

Juan Carlos Fernandez-Lopez

155. Center for Diabetes Research, University of Bergen, Bergen, Norway

Valeriya Lyssenko

156. Lund University Diabetes Center, Lund University, Malmo, Sweden

Valeriya Lyssenko

157. Genomic Research on Complex Diseases (GRC Group), CSIR–Centre for Cellular and Molecular Biology, Hyderabad, India

Meraj Ahmed & Giriraj R. Chandak

158. Epidemiology, School of Public Health, University of Alabama at Birmingham, Birmingham, AL, USA

Marguerite R. Irvin

159. Hunter Medical Research Institute, Newcastle, New South Wales, Australia

Christopher Oldmeadow, John Attia & Rodney J. Scott

160. Medical Research Institute, Kangbuk Samsung Hospital, Sungkyunkwan University School of Medicine, Seoul, Korea

Han-Na Kim

161. Department of Clinical Research Design and Evaluation, SAIHST, Sungkyunkwan University, Seoul, Korea

Han-Na Kim

162. Center for Cohort Studies, Total Healthcare Center, Kangbuk Samsung Hospital, Sungkyunkwan University School of Medicine, Seoul, Korea

Seungho Ryu

163. Department of Occupational and Environmental Medicine, Kangbuk Samsung Hospital, Sungkyunkwan University School of Medicine, Seoul, Korea

Seungho Ryu

164. Centre for Global Health Research, Usher Institute, University of Edinburgh, Edinburgh, UK

Paul R. H. J. Timmers, Katherine A. Kentistou, Harry Campbell, Peter K. Joshi & James F. Wilson

165. Thurston Arthritis Research Center, University of North Carolina, Chapel Hill, NC, USA

Liubov Arbeeva, Amanda E. Nelson & Yvonne M. Golightly

166. Division of Biomedical Informatics and Personalized Medicine, Department of Medicine, Anschutz Medical Campus, University of Colorado, Denver, Aurora, CO, USA

Leslie A. Lange

167. Ocular Epidemiology, Singapore Eye Research Institute, Singapore National Eye Centre, Singapore, Singapore

Xiaoran Chai, Ching-Yu Cheng, Tien-Yin Wong & Charumathi Sabanayagam

168. Department of Ophthalmology, National University of Singapore and National University Health System, Singapore, Singapore

Xiaoran Chai

169. Genomics and Molecular Medicine Unit, CSIR–Institute of Genomics and Integrative Biology, New Delhi, India

Gauri Prasad

170. Academy of Scientific and Innovative Research, CSIR–Institute of Genomics and Integrative Biology Campus, New Delhi, India

Gauri Prasad & Dwaipayan Bharadwaj

171. Departments of Ophthalmology and Human Genetics, Radboud University Nijmegen Medical Center, Nijmegen, The Netherlands

Laura Lorés-Motta, Marc Pauper & Anneke I. den Hollander

172. Vanderbilt Epidemiology Center, Division of Epidemiology, Vanderbilt University Medical Center, Nashville, TN, USA

Jirong Long, Qiuyin Cai, Xiao-Ou Shu & Wei Zheng

173. Department of Pediatrics, University of California, San Francisco, Oakland, CA,, USA

Elizabeth Theusch

174. National Center for Global Health and Medicine, Tokyo, Japan

Fumihiko Takeuchi & Norihiro Kato

175. Department of Genetics, University of North Carolina, Chapel Hill, NC, USA

Cassandra N. Spracklen & Karen L. Mohlke

176. Department of Biostatistics and Epidemiology, University of Massachusetts-Amherst, Amherst, MA,, USA

Cassandra N. Spracklen

177. Department of Cardiovascular Sciences, University of Leicester, Leicester, UK

Sophie C. Warner, Christopher P. Nelson, Peter S. Braund & Nilesh J. Samani

178. NIHR Leicester Biomedical Research Centre, Glenfield Hospital, Leicester, UK

Sophie C. Warner, Christopher P. Nelson, Peter S. Braund & Nilesh J. Samani

179. Beijing Institute of Ophthalmology, Beijing Key Laboratory of Ophthalmology and Visual Sciences, Beijing, China

Ya Xing Wang

180. Beijing Tongren Eye Center, Beijing Tongren Hospital, Capital Medical University, Beijing, China

Ya Xing Wang & Wen B. Wei

181. Institute of Genetics and Biophysics “Adriano Buzzati-Traverso”– CNR, Naples, Italy

Teresa Nutile, Daniela Ruggiero & Marina Ciullo

182. IRCCS Neuromed, Pozzilli, Isernia, Italy

Daniela Ruggiero & Marina Ciullo

183. Department of Psychiatry, Washington University, St Louis, MO, USA

Yun Ju Sung

184. Division of Endocrinology and Metabolism, Tri-Service General Hospital Songshan Branch, Taipei, Taiwan

Yi-Jen Hung

185. Rush Alzheimer's Disease Center, Rush University Medical Center, Chicago, IL, USA

Jingyun Yang & David A. Bennett

186. Department of Neurological Sciences, Rush University Medical Center, Chicago, IL, USA

Jingyun Yang & David A. Bennett

187. Department of Nephrology, University Hospital Regensburg, Regensburg, Germany

Mathias Gorski, Bettina Jung & Carsten A. Böger

188. Department of Medicine, Surgery and Health Sciences, University of Trieste, Trieste, Italy

Marco Brumat & Giorgia Girotto

189. Department of Molecular Epidemiology, German Institute of Human Nutrition Potsdam-Rehbruecke, Nuthetal, Germany

Karina Meidtner, Susanne Jäger & Matthias B. Schulze

190. German Center for Diabetes Research (DZD), Neuherberg, Germany

Karina Meidtner, Susanne Jäger, Matthias B. Schulze, Annette Peters & Christian Gieger

191. Department of Epidemiology, School of Public Health, University of Michigan, Ann Arbor, MI, USA

Lawrence F. Bielak, Jennifer A. Smith & Sharon L. R. Kardia

192. Department of Genetics and Bioinformatics, Dasman Diabetes Institute, Kuwait, Kuwait

Prashantha Hebbar, Fahd Al-Mulla & Thangavel Alphonse Thanaraj

193. Department of Nutrition and Dietetics, School of Health Science and Education, Harokopio University of Athens, Athens, Greece

Aliki-Eleni Farmaki & George Dedoussis

194. Department of Population Science and Experimental Medicine, University College London, London, UK

Aliki-Eleni Farmaki

195. Department of Neurology, Clinical Division of Neurogeriatrics, Medical University of Graz, Graz, Austria

Edith Hofer & Reinhold Schmidt

196. Institute for Medical Informatics, Statistics and Documentation, Medical University of Graz, Graz, Austria

Edith Hofer

197. Department of Bioinformatics and Genomics, University of North Carolina at Charlotte, Charlotte, NC, USA

Maoxuan Lin

198. Institute for Maternal and Child Health, IRCCS Burlo Garofolo, Trieste, Italy

Maria Pina Concas

199. Institute of Genetic and Biomedical Research, National Research Council of Italy, UOS of Sassari, Sassari, Italy

Simona Vaccargiu

200. Department of Epidemiology, University Medical Center Groningen, University of Groningen, Groningen, The Netherlands

Peter J. van der Most & Harold Snieder

201. Research Centre of Applied and Preventive Cardiovascular Medicine, University of Turku, Turku, Finland

Niina Pitkänen, Olli T. Raitakari & Katja Pahkala

202. Centre for Population Health Research, University of Turku and Turku University Hospital, Turku, Finland

Niina Pitkänen, Olli T. Raitakari & Katja Pahkala

203. Sleep Medicine and Circadian Disorders, Brigham and Women's Hospital, Boston, MA, USA

Brian E. Cade, Jiwon Lee & Susan Redline

204. Division of Sleep Medicine, Harvard Medical School, Boston, MA, USA

Brian E. Cade & Susan Redline

205. Central Diagnostics Laboratory, Division Laboratories, Pharmacy, and Biomedical Genetics, University Medical Center Utrecht, Utrecht University, Utrecht, The Netherlands

Sander W. van der Laan, Gerard Pasterkamp & Imo E. Hoefer

206. Laboratory of Epidemiology and Population Science, National Institute on Aging Intramural Research Program, NIH Biomedical Research Center, NIA, Baltimore, MD, USA

Kumaraswamy Naidu Chitrala

207. Department of Functional Genomics, Interfaculty Institute for Genetics and Functional Genomics, University of Greifswald and University Medicine Greifswald, Greifswald, Germany

Stefan Weiss & Georg Homuth

208. Oneomics, Bucheon, Korea

Jong Young Lee

209. Department of Biomedical Science, Hallym University, Chuncheon, Korea

Hyeok Sun Choi, Jae Hun Shin & Yoon Shin Cho

210. Centre for Bone and Arthritis Research, Department of Internal Medicine and Clinical Nutrition, Institute of Medicine, Sahlgrenska Academy, University of Gothenburg, Gothenburg, Sweden

Maria Nethander, Claes Ohlsson & Dan Mellström

211. Bioinformatics Core Facility, Sahlgrenska Academy, University of Gothenburg, Gothenburg, Sweden

Maria Nethander

212. Institute of Medical Informatics and Statistics, Kiel University, Kiel, Germany

Sandra Freitag-Wolf

213. Institute of Translational Genomics, Helmholtz Zentrum München–German Research Center for Environmental Health, Neuherberg, Germany

Lorraine Southam, Nigel W. Rayner & Eleftheria Zeggini

214. Wellcome Trust Sanger Institute, Hinxton, UK

Lorraine Southam & Eleftheria Zeggini

215. Oxford Centre for Diabetes Endocrinology and Metabolism, Oxford, UK

Nigel W. Rayner

216. School of Medicine and Public Health, College of Health, Medicine and Wellbeing, University of Newcastle, Newcastle, New South Wales, Australia

Carol A. Wang & Craig E. Pennell

217. Center for Geriatrics and Gerontology, Division of Endocrinology and Metabolism, Department of Internal Medicine, Taichung Veterans General Hospital, Taichung, Taiwan

Shih-Yi Lin

218. School of Medicine, National Yang-Ming University, Taipei, Taiwan

Shih-Yi Lin

219. School of Medicine, National Defense Medical Center, Taipei, Taiwan

Shih-Yi Lin

220. Division of Endocrinology and Metabolism, Department of Internal Medicine, Taichung Veterans General Hospital, Taichung, Taiwan

Jun-Sing Wang

221. Department of Medicine, School of Medicine, National Yang-Ming University, Taipei, Taiwan

Jun-Sing Wang

222. Department of Kinesiology, Université Laval, Quebec, Quebec, Canada

Christian Couture & Louis Pérusse

223. Department of Clinical Chemistry, Fimlab Laboratories, Tampere, Finland

Leo-Pekka Lyytikäinen & Terho Lehtimäki

224. Department of Clinical Chemistry, Finnish Cardiovascular Research Center–Tampere, Faculty of Medicine and Health Technology, Tampere University, Tampere, Finland

Leo-Pekka Lyytikäinen & Terho Lehtimäki

225. Department of Cardiology, Heart Center, Tampere University Hospital, Tampere, Finland

Kjell Nikus

226. Department of Cardiology, Finnish Cardiovascular Research Center–Tampere, Faculty of Medicine and Health Technology, Tampere University, Tampere, Finland

Kjell Nikus

227. University of Queensland Diamantina Institute, Translational Research Institute, Brisbane, Queensland, Australia

Gabriel Cuellar-Partida

228. Department of Medicine, Bornholms Hospital, Ronne, Denmark

Henrik Vestergaard

229. School of Public Health, University of Alabama at Birmingham, Birmingham, AL, USA

Bertha Hildalgo

230. Cardiology, Division Heart and Lungs, University Medical Center Utrecht, Utrecht University, Utrecht, The Netherlands

Jessica van Setten, Folkert W. Asselbergs & Adriaan O. Kraaijeveld

231. Department of Population and Quantitative Health Sciences, Case Western Reserve University, Cleveland, OH, USA

Xiaoyin Li & Xiaofeng Zhu

232. Division of Endocrinology, Ohio State University, Columbus, OH, USA

Rebecca D. Jackson

233. Department of Epidemiology, University of Washington, Seattle, WA, USA

Alexander P. Reiner

234. School of Medicine and Health Sciences, George Washington University, Washington, DC, USA

Lisa Warsinger Martin

235. Department of Epidemiology, School of Public Health, Peking University Health Science Center, Beijing, China

Liming Li

236. Institute for Laboratory Medicine, University Hospital Leipzig, Leipzig, Germany

Joachim Thiery

237. Laboratory of Epidemiology and Population Sciences, National Institute on Aging, NIH, Baltimore, MD, USA

Lenore J. Launer

238. Department of Clinical Physiology and Nuclear Medicine, Turku University Hospital, Turku, Finland

Olli T. Raitakari

239. Department of Environmental and Preventive Medicine, Jichi Medical University School of Medicine, Shimotsuke, Japan

Sahoko Ichihara

240. Centre for Population Health Sciences, Usher Institute, University of Edinburgh, Edinburgh, UK

Sarah H. Wild

241. Department of Functional Pathology, Shimane University School of Medicine, Izumo, Japan

Toru Nabika

242. Department of Pediatrics and Adolescent Medicine, Turku University Hospital and University of Turku, Turku, Finland

Harri Niinikoski

243. Department of Physiology, University of Turku, Turku, Finland

Harri Niinikoski

244. Faculty of Medicine, University of Split, Split, Croatia

Ivana Kolcic & Ozren Polasek

245. Medical Department III–Endocrinology, Nephrology, Rheumatology, University of Leipzig Medical Center, Leipzig, Germany

Peter Kovacs & Anke Tönjes

246. Department of Nutrition–Dietetics, Harokopio University, Eleftheriou Venizelou, Athens, Greece

Tota Giardoglou

247. Department of Clinical Gene Therapy, Osaka University Graduate School of Medicine, Suita, Japan

Tomohiro Katsuya

248. Department of Geriatric and General Medicine, Osaka University Graduate School of Medicine, Suita, Japan

Tomohiro Katsuya

249. Department of Vascular Surgery, Division of Surgical Specialties, University Medical Center Utrecht, Utrecht University, Utrecht, The Netherlands

Dominique de Kleijn & Gert J. de Borst

250. Corneal Dystrophy Research Institute, Department of Ophthalmology, Yonsei University College of Medicine, Seoul, Korea

Eung Kweon Kim

251. Department of Radiology and Nuclear Medicine, Erasmus MC, University Medical Center Rotterdam, Rotterdam, The Netherlands

Hieab H. H. Adams

252. Julius Centre for Health Sciences and Primary Care, University Medical Centre Utrecht, Utrecht, The Netherlands

Joline W. J. Beulens

253. Second Department of Cardiology, Medical School, National and Kapodistrian University of Athens, Attikon University Hospital, Athens, Greece

Loukianos S. Rallidis

254. Center for Vision Research, Department of Ophthalmology and The Westmead Institute, University of Sydney, Sydney, New South Wales, Australia

Paul Mitchell

255. Menzies Institute for Medical Research, School of Medicine, University of Tasmania, Hobart, Tasmania, Australia

Alex W. Hewitt

256. Centre for Eye Research Australia, University of Melbourne, Melbourne, Victoria, Australia

Alex W. Hewitt

257. Department of Clinical Physiology, Tampere University Hospital, Tampere, Finland

Mika Kähönen

258. Department of Clinical Physiology, Finnish Cardiovascular Research Center–Tampere, Faculty of Medicine and Health Technology, Tampere University, Tampere, Finland

Mika Kähönen

259. Institute of Nutrition and Functional Foods (INAF), Université Laval, Quebec, Quebec, Canada

Louis Pérusse

260. Pennington Biomedical Research Center, Baton Rouge, LA, USA

Claude Bouchard

261. Medical School, Faculty of Health and Medical Sciences, University of Western Australia, Perth, Western Australia, Australia

Trevor A. Mori

262. Institute of Epidemiology, Kiel University, Kiel, Germany

Wolfgang Lieb

263. Institute of Clinical Molecular Biology, Kiel University, Kiel, Germany

Andre Franke

264. Department of Drug Treatment, Sahlgrenska University Hospital, Gothenburg, Sweden

Claes Ohlsson

265. Geriatric Medicine, Institute of Medicine, Sahlgrenska Academy, University of Gothenburg, Gothenburg, Sweden

Dan Mellström

266. Department of Internal Medicine, Ewha Womans University School of Medicine, Seoul, Korea

Hyejin Lee

267. Division of Cancer Control and Population Sciences, UPMC Hillman Cancer Center, University of Pittsburgh, Pittsburgh, PA, USA

Jian-Min Yuan

268. Department of Epidemiology, Graduate School of Public Health, University of Pittsburgh, Pittsburgh, PA, USA

Jian-Min Yuan

269. Healthy Longevity Translational Research Programme, Yong Loo Lin School of Medicine, National University of Singapore, Singapore, Singapore

Woon-Puay Koh

270. Singapore Institute for Clinical Sciences, Agency for Science Technology and Research (A*STAR), Singapore, Singapore

Woon-Puay Koh

271. Department of Endocrinology and Metabolism, Kyung Hee University School of Medicine, Seoul, Korea

Sang Youl Rhee & Jeong-Taek Woo

272. Institute for Community Medicine, University Medicine Greifswald, Greifswald, Germany

Henry Völzke

273. Laboratory of Epidemiology and Population Science, National Institute on Aging Intramural Research Program, NIH Biomedical Research Center, Baltimore, MD, USA

Michele K. Evans & Alan B. Zonderman

274. Paavo Nurmi Centre, Sports and Exercise Medicine Unit, Department of Physical Activity and Health, University of Turku, Turku, Finland

Katja Pahkala

275. Interdisciplinary Center Psychopathology and Emotion Regulation (ICPE), University of Groningen, University Medical Center

Groningen, Groningen, The Netherlands

Albertine J. Oldehinkel

276. Institute of Molecular Genetics, National Research Council of Italy, Pavia, Italy

Ginevra Biino

277. Gottfried Schatz Research Center for Cell Signaling, Metabolism and Aging, Medical University of Graz, Graz, Austria

Helena Schmidt

278. Local Health Unit Toscana Centro, Firenze, Italy

Stefania Bandinelli

279. Institute of Nutritional Science, University of Potsdam, Nuthetal, Germany

Matthias B. Schulze

280. Institute for Maternal and Child Health IRCCS "Burlo Garofolo", Trieste, Italy

Giorgia Girotto

281. Department of Nephrology, Diabetology, Rheumatology, Traunstein Hospital, Traunstein, Germany

Carsten A. Böger

282. KfH Kidney Center Traunstein, Traunstein, Germany

Carsten A. Böger

283. Center for Translational and Systems Neuroimmunology, Department of Neurology, Columbia University Medical Center, New York, NY,

USA

Philip L. De Jager

284. Program in Medical and Population Genetics, Broad Institute, Cambridge, MA, USA

Philip L. De Jager

285. Medical School, National and Kapodistrian University Athens, Athens, Greece

Vasiliki Mamakou

286. Dromokaiteio Psychiatric Hospital, Athens, Greece

Vasiliki Mamakou

287. Clinical Pharmacology, William Harvey Research Institute, Queen Mary University of London, London, UK

Morris Brown

288. Department of Ophthalmology, Medical Faculty Mannheim, Heidelberg University, Mannheim, Germany

Jost B. Jonas

289. Beijing Institute of Ophthalmology, Beijing Key Laboratory of Ophthalmology and Visual Sciences, Beijing Tongren Eye Center, Beijing Tongren Hospital, Capital Medical University, Beijing, China

Jost B. Jonas

290. Institute of Molecular and Clinical Ophthalmology Basel, Basel, Switzerland

Jost B. Jonas

291. Harvard Medical School, Boston, MA, USA

Daniel I. Chasman

292. Harvard Medical School, Boston, MA, USA

Josep M. Mercader, Julie E. Buring & Paul M. Ridker

293. Department of Human Genetics, David Geffen School of Medicine at UCLA, University of California, Los Angeles, Los Angeles, CA, USA

Päivi Pajukanta

294. Unidad de Biología Molecular y Medicina Genómica, Instituto Nacional de Ciencias Médicas y Nutrición Salvador Zubirán, Mexico, Mexico

Teresa Tusié-Luna

295. Instituto de Investigaciones Biomédicas, UNAM, Ciudad de México, Mexico, Mexico

Teresa Tusié-Luna

296. Unidad de Biología Molecular y Medicina Genómica, Instituto de Investigaciones Bimédicas UNAM/ Instituto Nacional de Ciencias Médicas y Nutrición Salvador Zubirán, Mexico, Mexico

Teresa Tusié-Luna

297. Departamento de Endocrinología y Metabolismo, Instituto Nacional de Ciencias Médicas y Nutrición Salvador Zubirán, Mexico, Mexico

Carlos A. Aguilar-Salinas

298. Dirección de Nutrición and Unidad de Estudios de Enfermedades Metabólicas, Instituto Nacional de Ciencias Médicas y Nutrición Salvador Zubirán, Mexico, Mexico

Carlos A. Aguilar-Salinas

299. Department of Nutrition, Gillings School of Global Public Health, University of North Carolina, Chapel Hill, NC, USA

Linda S. Adair & Penny Gordon-Larsen

300. Carolina Population Center, University of North Carolina, Chapel Hill, NC, USA

Linda S. Adair & Penny Gordon-Larsen

301. USC–Office of Population Studies Foundation, University of San Carlos, Cebu City, Philippines

Sonny Augustin Bechayda

302. Department of Anthropology, Sociology, and History, University of San Carlos, Cebu City, Philippines

Sonny Augustin Bechayda

303. Department of Medicine, Faculty of Medicine, University of Kelaniya, Ragama, Sri Lanka

H. Janaka de Silva

304. Department of Public Health, Faculty of Medicine, University of Kelaniya, Ragama, Sri Lanka

Ananda R. Wickremasinghe

305. Departments of Pediatrics and Medicine, University of California, San Francisco, San Francisco, CA, USA

Ronald M. Krauss

306. Institute of Biomedical Sciences, Academia Sinica, Taipei, Taiwan

Jer-Yuarn Wu

307. Systems Genomics Laboratory, School of Biotechnology, Jawaharlal Nehru University, New Delhi, India

Dwaipayan Bharadwaj

308. Department of Medicine, University of Mississippi Medical Center, Jackson, MS, USA

Adolfo Correa

309. Department of Physiology and Biophysics, University of Mississippi Medical Center, Jackson, MS, USA

James G. Wilson

310. Department of Medical Sciences, Uppsala University, Uppsala, Sweden

Lars Lind

311. Department of Paediatrics, Yong Loo Lin School of Medicine, National University of Singapore, Singapore, Singapore

Chew-Kiat Heng

312. Khoo Teck Puat–National University Children's Medical Institute, National University Health System, Singapore, Singapore

Chew-Kiat Heng

313. Department of Medicine, University of North Carolina, Chapel Hill, NC, USA

Amanda E. Nelson

314. Department of Epidemiology, Gillings School of Global Public Health, University of North Carolina, Chapel Hill, NC, USA

Yvonne M. Golightly

315. Injury Prevention Research Center, University of North Carolina, Chapel Hill, NC, USA

Yvonne M. Golightly

316. Division of Physical Therapy, University of North Carolina, Chapel Hill, NC, USA

Yvonne M. Golightly

317. Department of Psychiatry, Amsterdam UMC, Vrije Universiteit Amsterdam, Amsterdam, The Netherlands

Brenda Penninx

318. Amsterdam Public Health Research Institute, VU Medical Center Amsterdam, Amsterdam, The Netherlands

Brenda Penninx & Eco J. C. de Geus

319. Department of Biochemistry, College of Medicine, Ewha Womans University, Seoul, Korea

Hyung-Lae Kim

320. Faculty of Health and Medicine, University of Newcastle, Newcastle, New South Wales, Australia

John Attia & Rodney J. Scott

321. Division of Biostatistics, Washington University School of Medicine, St Louis, MO, USA

D. C. Rao

322. University of Kentucky, College of Public Health, Lexington, KY, USA

Donna K. Arnett

323. Institute of Cellular Medicine (Diabetes), The Medical School, Newcastle University, Newcastle upon Tyne, UK

Mark Walker

324. Department of Population Health, Finnish Institute for Health and Welfare, Helsinki, Finland

Heikki A. Koistinen

325. University of Helsinki and Department of Medicine, Helsinki University Hospital, Helsinki, Finland

Heikki A. Koistinen

326. Minerva Foundation Institute for Medical Research, Helsinki, Finland

Heikki A. Koistinen

327. Academy of Scientific and Innovative Research (AcSIR), New Delhi, India

Giriraj R. Chandak

328. Diabetology Research Centre, KEM Hospital and Research Centre, Pune, India

Chittaranjan S. Yajnik

329. Programs in Metabolism and Medical and Population Genetics, Broad Institute of MIT and Harvard, Cambridge, MA, USA

Josep M. Mercader

330. Diabetes Unit and Center for Genomic Medicine, Massachusetts General Hospital, Boston, MA, USA

Josep M. Mercader

331. Instituto Nacional de Salud Publica y Centro de Estudios en Diabetes, Cuernavaca, Mexico

Clicerio Gonzalez Villalpando

332. Instituto Nacional de Medicina Genómica, Mexico, Mexico

Lorena Orozco

333. Human Genetics Center, School of Public Health, University of Texas Health Science Center at Houston, Houston, TX, USA

Myriam Fornage

334. Yong Loo Lin School of Medicine, National University of Singapore and National University Health System, Singapore, Singapore

E. Shyong Tai & Rob M. van Dam

335. Kurume University School of Medicine, Kurume, Japan

Mitsuhiro Yokota

336. Genetics, Merck Sharp & Dohme, Kenilworth, NJ, USA

Dermot F. Reilly

337. Oxford Centre for Diabetes, Endocrinology and Metabolism, University of Oxford, Oxford, UK

Mark I. McCarthy

338. Population Health and Genomics, University of Dundee, Ninwells Hospital and Medical School, Dundee, UK

Colin N. A. Palmer

339. Intramural Research Program, National Institute on Aging, Baltimore, MD, USA

Eleanor Simonsick

340. The Eye Hospital, School of Ophthalmology and Optometry, Wenzhou Medical University, Wenzhou, China

Fan Lu & Jia Qu

341. Shanghai Institute of Nutrition and Health, University of Chinese Academy of Sciences, Chinese Academy of Sciences, Shanghai, China

Shaofeng Huo, Huaixing Li & Xu Lin

342. Synlab Academy, Synlab, Mannheim, Germany

Winfried März

343. Clinical Institute of Medical and Chemical Laboratory Diagnostics, Medical University of Graz, Graz, Austria

Winfried März

344. Unidad de Investigacion Medica en Bioquimica, Hospital de Especialidades, Centro Medico Nacional Siglo XXI, Instituto Mexicano del Seguro Social, Mexico, Mexico

Miguel Cruz

345. Faculty of Medicine, University of Iceland, Reykjavik, Iceland

Vilmundur Gudnason, Unnur Thorsteinsdottir & Kari Stefansson

346. Department of Medicine, Faculty of Medecine, Université de Montréal, Quebec, Quebec, Canada

Jean-Claude Tardif

347. Department of Medicine, Faculty of Medicine, Université de Montréal, Montreal, Quebec, Canada

Guillaume Lettre

348. Department of Biomedical Data Sciences, Section Molecular Epidemiology, Leiden University Medical Center, Leiden, The Netherlands

Leen M. ’t Hart

349. Amsterdam UMC, Department of General Practice and Elderly Care, Amsterdam Public Health Research Institute, Amsterdam, The Netherlands

Petra J. M. Elders

350. Department of Genetics, University of Pennsylvania, Philadelphia, PA, USA

Struan F. A. Grant & Daniel J. Rader

351. Department of Medicine, Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA, USA

Daniel J. Rader & Kyong-Mi Chang

352. Department of Surgery, University of Pennsylvania, Philadelphia, PA, USA

Scott M. Damrauer

353. Corporal Michael Crescenz VA Medical Center, Philadelphia, PA, USA

Scott M. Damrauer

354. Institute of Social and Economic Research, University of Essex, Essex, UK

Meena Kumari

355. The Mindich Child Health and Development Institute, Icahn School of Medicine at Mount Sinai, New York, NY, USA

Ruth J. F. Loos

356. Cardiovascular Health Research Unit, Departments of Medicine, Epidemiology and Health Services, University of Washington, Seattle, WA, USA

Bruce M. Psaty

357. Kaiser Permanent Washington Health Research Institute, Seattle, WA, USA

Bruce M. Psaty

358. Institute of Regional Health Research, University of Southern Denmark, Odense, Denmark

Ivan Brandslund

359. Danish Aging Research Center, University of Southern Denmark, Odense, Denmark

Kaare Christensen

360. Public Health, Faculty of Medicine, University of Helsinki, Helsinki, Finland

Samuli Ripatti

361. Broad Institute of MIT and Harvard, Cambridge, MA, USA

Samuli Ripatti

362. Center for Applied Genomics, Children's Hospital of Philadelphia, Philadelphia, PA, USA

Hakon Hakonarson

363. Department of Pediatrics, Perelman School of Medicine, The University of Pennsylvania, Philadelphia, PA, USA

Hakon Hakonarson & Struan F. A. Grant

364. Division of Human Genetics, Children's Hospital of Philadelphia, Philadelphia, PA, USA

Struan F. A. Grant

365. Institute of Genetic Epidemiology, Department of Genetics and Pharmacology, Medical University of Innsbruck, Innsbruck, Austria

Florian Kronenberg

366. School of Medicine, Southern University of Science and Technology, Shenzhen, China

Dongfeng Gu

367. Institute for Cardiogenetics, University of Lübeck, DZHK (German Research Centre for Cardiovascular Research), Partner site Hamburg/Lübeck/Kiel, and University Heart Center Lübeck, Lübeck, Germany

Jeanette Erdmann

368. Deutsches Herzzentrum München, Klinik für Herz- und Kreislauferkrankungen, Technische Universität München, Munich, Germany

Heribert Schunkert

369. Deutsches Zentrum für Herz-Kreislauf-Forschung (DZHK) e.V., Partner site Munich Heart Alliance, Munich, Germany

Heribert Schunkert

370. Department of Clinical Medicine, Faculty of Health and Medical Sciences, University of Copenhagen, Copenhagen, Denmark

Allan Linneberg

371. Netherlands Heart Institute, Utrecht, The Netherlands

J. Wouter Jukema

372. Division of Cardiology, Department of Medicine, Massachusetts General Hospital, Boston, MA, USA

Amit V. Khera

373. Program of Medical and Population Genetics, Broad Institute, Cambridge, MA, USA

Amit V. Khera & Sekar Kathiresan

374. Center for Genomic Medicine, Massachusetts General Hospital, Boston, MA, USA

Amit V. Khera

375. Department of Medicine, Harvard Medical School, Boston, MA, USA

Amit V. Khera & Sekar Kathiresan

376. Northern Finland Birth Cohorts, Infrastructure for Population Studies, Faculty of Medicine, University of Oulu, Oulu, Finland

Minna Männikkö

377. Center for Life Course Health Research, Faculty of Medicine, University of Oulu, Oulu, Finland

Marjo-Riitta Jarvelin

378. Biocenter of Oulu, University of Oulu, Oulu, Finland

Marjo-Riitta Jarvelin

379. University Center for Primary Care and Public Health, Lausanne, Switzerland

Zoltan Kutalik

380. Institute for Genetic and Biomedical Research, Italian National Council of Research (IRGB CNR), Cagliari, Italy

Francesco Cucca

381. University of Sassari, Sassari, Italy

Francesco Cucca

382. Department of Clinical Epidemiology, Leiden University Medical Center, Leiden, The Netherlands

Dennis O. Mook-Kanamori

383. Department of Public Health and Primary Care, Leiden University Medical Center, Leiden, The Netherlands

Dennis O. Mook-Kanamori

384. Department of Internal Medicine, Division of Endocrinology, Leiden University Medical Center, Leiden, the Netherlands

Ko Willems van Dijk

385. Einthoven Laboratory for Experimental Vascular Medicine, Leiden University Medical Center, Leiden, the Netherlands

Ko Willems van Dijk

386. Department of Human Genetics, Leiden University Medical Center, Leiden, The Netherlands

Ko Willems van Dijk

387. Division of Cardiovascular Medicine, Radcliffe Department of Medicine, John Radcliffe Hospital, University of Oxford, Oxford, UK

Hugh Watkins

388. Population Health Research Institute, St George's University of London, London, UK

David P. Strachan

389. National Heart and Lung Institute, Imperial College London, London, UK

Peter Sever

390. School of Public Health, Imperial College London, London, UK

Neil Poulter

391. OCDEM, University of Oxford, Churchill Hospital, Oxford, UK

Fredrik Karpe & Matt J. Neville

392. NIHR Oxford Biomedical Research Centre, Churchill Hospital, Oxford, UK

Fredrik Karpe & Matt J. Neville

393. Ophthalmology and Visual Sciences Academic Clinical Program (Eye ACP), Duke–NUS Medical School, Singapore, Singapore

Ching-Yu Cheng, Tien-Yin Wong & Charumathi Sabanayagam

394. DZHK (German Centre for Cardiovascular Research), Munich Heart Alliance Partner Site, Munich, Germany

Annette Peters

395. University of Exeter Medical School, University of Exeter, Exeter, UK

Andrew T. Hattersley

396. Department of Medical Epidemiology and Biostatistics, Karolinska Institutet, Stockholm, Sweden

Nancy L. Pedersen & Patrik K. E. Magnusson

397. Amsterdam Public Health, VU Medical Center Amsterdam, Amsterdam, The Netherlands

Dorret I. Boomsma

398. Amsterdam Reproduction and Development Research Institute, VU Medical Center Amsterdam, Amsterdam, the Netherlands

Dorret I. Boomsma

399. Framingham Heart Study, National Heart, Lung, and Blood Institute, US National Institutes of Health, Bethesda, MD, USA

L. Adrienne Cupples

400. Department of Genetics, School of Medicine, Mashhad University of Medical Sciences, Mashhad, Iran

Mohsen Ghanbari

401. Department of Genetics, Shanghai-MOST Key Laboratory of Health and Disease Genomics, Chinese National Human Genome Center at Shanghai, Shanghai, China

Wei Huang

402. TUM School of Medicine, Technical University of Munich (TUM) and Klinikum Rechts der Isar, Munich, Germany

Eleftheria Zeggini

403. Institute of Clinical Medicine, Internal Medicine, University of Eastern Finland and Kuopio University Hospital, Kuopio, Finland

Johanna Kuusisto & Markku Laakso

404. Stanford Cardiovascular Institute, Stanford University, Stanford, CA, USA

Erik Ingelsson

405. Stanford Diabetes Research Center, Stanford University, Stanford, CA, USA

Erik Ingelsson

406. Regeneron Pharmaceuticals, Tarrytown, NY, USA

Goncalo Abecasis & Aris Baras

407. Lee Kong Chian School of Medicine, Nanyang Technological University, Singapore, Singapore

John C. Chambers

408. Imperial College Healthcare NHS Trust, Imperial College London, London, UK

John C. Chambers

409. MRC–PHE Centre for Environment and Health, Imperial College London, London, UK

Jaspal S. Kooner

410. National Heart and Lung Institute, Imperial College London, London, UK

Jaspal S. Kooner

411. Institute for Minority Health Research, University of Illinois College of Medicine, Chicago, IL, USA

Martha Daviglus

412. Department of Biostatistics, Harvard T. H. Chan School of Public Health, Boston, MA, USA

Peter Kraft

413. QIMR Berghofer Medical Research Institute, Brisbane, Queensland, Australia

Nicholas G. Martin & John B. Whitfield

414. Faisalabad Institute of Cardiology, Faislabad, Pakistan

Shahid Abbas

415. Department of Medicine, Columbia University Irving Medical Center, New York, NY, USA

Danish Saleheen

416. Department of Cardiology, Columbia University Irving Medical Center, New York, NY, USA

Danish Saleheen

417. Big Data Institute, University of Oxford, Oxford, UK

Robin G. Walters

418. National Institute for Health Research Oxford Biomedical Research Centre, Oxford University Hospitals, Oxford, UK

Michael V. Holmes

419. Aberdeen Centre for Health Data Science, School of Medicine, Medical Science and Nutrition, University of Aberdeen, Aberdeen, UK

Corri Black

420. Division of Population Health and Genomics, Ninewells Hospital and Medical School, University of Dundee, Dundee, UK

Blair H. Smith

421. Biomedical and Translational Informatics, Geisinger Health, Danville, PA, USA

Anne E. Justice

422. Department of Biomedical Informatics, Vanderbilt University Medical Center, Nashville, TN, USA

Wei-Qi Wei

423. Departments of Medicine (Medical Genetics) and Genome Sciences, University of Washington Medical Center, Seattle, WA, USA

Gail P. Jarvik

424. Center for Autoimmune Genomics and Etiology, Cincinnati Children's Hospital Medical Center (CCHMC), Cincinnati, OH, USA

Bahram Namjou

425. Division of Endocrinology, Metabolism, and Molecular Medicine, Department of Medicine, Northwestern University, Feinberg School of Medicine, Chicago, IL, USA

M. Geoffrey Hayes

426. Department of Anthropology, Northwestern University, Evanston, IL, USA

M. Geoffrey Hayes

427. Center for Genetic Medicine, Northwestern University, Feinberg School of Medicine, Chicago, IL, USA

M. Geoffrey Hayes

428. HUNT Research Centre, Department of Public Health and Nursing, NTNU, Norwegian University of Science and Technology, Levanger, Norway

Kristian Hveem & Bjørn Olav Åsvold

429. Department of Medicine, Levanger Hospital, Nord-Trøndelag Hospital Trust, Levanger, Norway

Kristian Hveem

430. Department of Endocrinology, St Olavs Hospital, Trondheim University Hospital, Trondheim, Norway

Bjørn Olav Åsvold

431. RIKEN Center for Integrative Medical Sciences, Yokohama, Japan

Michiaki Kubo

432. Laboratory of Complex Trait Genomics, Department of Computational Biology and Medical Sciences, Graduate School of Frontier Sciences, The University of Tokyo, Tokyo, Japan

Yoichiro Kamatani

433. Laboratory of Statistical Immunology, WPI Immunology Frontier Research Center, Osaka University, Osaka, Japan

Yukinori Okada

434. Integrated Frontier Research for Medical Science Division, Institute for Open and Transdisciplinary Research Initiatives, Osaka University, Osaka, Japan

Yukinori Okada

435. Division of Molecular Pathology, Institute of Medical Science, The University of Tokyo, Tokyo, Japan

Yoshinori Murakami

436. VA Boston Healthcare System, Boston, MA, USA

Yuk-Lam Ho, Kelly Cho, Christopher J. O'Donnell & John M. Gaziano

437. VA Informatics and Computing Infrastructure, VA Salt Lake City Health Care System, Salt Lake City, UT, USA

Julie A. Lynch

438. University of Massachusetts, Boston, MA, USA

Julie A. Lynch

439. Cardiovascular Institute, Stanford University School of Medicine, Stanford, CA, USA

Philip S. Tsao & Themistocles L. Assimes

440. Corporal Michael J. Crescenz VA Medical Center, Philadelphia, PA, USA

Kyong-Mi Chang

441. Department of Medicine, Brigham Women's Hospital, Boston, MA, USA

Kelly Cho, Christopher J. O'Donnell & John M. Gaziano

442. Atlanta VA Medical Center, Atlanta, GA, USA

Peter Wilson

443. Division of Cardiology, Emory University School of Medicine, Atlanta, GA, USA

Peter Wilson

444. School of Electrical and Information Engineering, University of the Witwatersrand, Johannesburg, South Africa

Scott Hazelhurst

445. Division of Human Genetics, National Health Laboratory Service and School of Pathology, Faculty of Health Sciences, University of the Witwatersrand, Johannesburg, South Africa

Michèle Ramsay

446. School of Basic and Medical Biosciences, Faculty of Life Sciences and Medicine, King's College London, London, UK

Richard C. Trembath

447. Departments of Pediatrics and Genetics, Harvard Medical School, Boston, MA, USA

Joel N. Hirschhorn

448. Center for Genomic Medicine, Department of Medicine, Massachusetts General Hospital, Boston, MA, USA

Sekar Kathiresan

449. Cardiology Division, Massachusetts General Hospital, Harvard Medical School, Boston, MA, USA

Pradeep Natarajan

450. Department of Medicine, Massachusetts General Hospital, Harvard Medical School, Boston, MA, USA

Pradeep Natarajan

451. Cardiovascular Research Center and Center for Genomic Medicine, Massachusetts General Hospital, Boston, MA, USA

Pradeep Natarajan

452. Centre for Genetics and Genomics Versus Arthritis, Centre for Musculoskeletal Research, Division of Musculoskeletal and Dermatological Sciences, The University of Manchester, Manchester, UK

Andrew P. Morris

453. Princess Al-Jawhara Al-Brahim Centre of Excellence in Research of Hereditary Disorders (PACER-HD), King Abdulaziz University, Jeddah, Saudi Arabia

Panos Deloukas

454. Department of Human Genetics, University of Michigan, Ann Arbor, MI, USA

Cristen J. Willer

455. Department of Genetic Medicine, Weill Cornell Medicine-Qatar, Doha, Qatar

Noha A. Yousri & Steven C. Hunt

456. Department of Computer and Systems Engineering, Alexandria University, Alexandria, Egypt

Noha A. Yousri

457. Department of Clinical Genetics, Erasmus MC, University Medical Center Rotterdam, Rotterdam, The Netherlands

Hieab H. H. Adams

458. Department of Internal Medicine, University of Utah, Salt Lake City, UT, USA

Steven C. Hunt

Consortia

VA Million Veteran Program

- Shoa L. Clarke
- , Qin Hui
- , Derek Klarin
- , Xiang Zhu
- , Scott M. Damrauer
- , Yuk-Lam Ho
- , Julie A. Lynch
- , Daniel J. Rader
- , Philip S. Tsao
- , Kyong-Mi Chang
- , Kelly Cho
- , Christopher J. O’Donnell
- , John M. Gaziano
- , Peter Wilson
- , Themistocles L. Assimes
- , Yan V. Sun
- & Ioanna Ntalla

Global Lipids Genetics Consortium

- Sarah E. Graham
- , Shoa L. Clarke
- , Kuan-Han H. Wu

- , Stavroula Kanoni
- , Greg J. M. Zajac
- , Shweta Ramdas
- , Ida Surakka
- , Ioanna Ntalla
- , Sailaja Vedantam
- , Thomas W. Winkler
- , Adam E. Locke
- , Eirini Marouli
- , Mi Yeong Hwang
- , Sohee Han
- , Akira Narita
- , Ananyo Choudhury
- , Amy R. Bentley
- , Kenneth Ekoru
- , Anurag Verma
- , Bhavi Trivedi
- , Hilary C. Martin
- , Karen A. Hunt
- , Qin Hui
- , Derek Klarin
- , Xiang Zhu
- , Gudmar Thorleifsson
- , Anna Helgadottir
- , Daniel F. Gudbjartsson
- , Hilma Holm
- , Isleifur Olafsson
- , Masato Akiyama
- , Saori Sakaue
- , Chikashi Terao
- , Masahiro Kanai
- , Wei Zhou
- , Ben M. Brumpton
- , Humaira Rasheed
- , Sanni E. Ruotsalainen
- , Aki S. Havulinna
- , Yogasudha Veturi

- , QiPing Feng
- , Elisabeth A. Rosenthal
- , Todd Lingren
- , Jennifer Allen Pacheco
- , Sarah A. Pendergrass
- , Jeffrey Haessler
- , Franco Giulianini
- , Yuki Bradford
- , Jason E. Miller
- , Archie Campbell
- , Kuang Lin
- , Iona Y. Millwood
- , George Hindy
- , Asif Rasheed
- , Jessica D. Faul
- , Wei Zhao
- , David R. Weir
- , Constance Turman
- , Hongyan Huang
- , Mariaelisa Graff
- , Anubha Mahajan
- , Michael R. Brown
- , Weihua Zhang
- , Ketian Yu
- , Ellen M. Schmidt
- , Anita Pandit
- , Stefan Gustafsson
- , Xianyong Yin
- , Jian'an Luan
- , Jing-Hua Zhao
- , Fumihiko Matsuda
- , Hye-Mi Jang
- , Kyungheon Yoon
- , Carolina Medina-Gomez
- , Achilleas Pitsillides
- , Jouke Jan Hottenga
- , Gonneke Willemsen

- , Andrew R. Wood
- , Yingji Ji
- , Zishan Gao
- , Simon Haworth
- , Ruth E. Mitchell
- , Jin Fang Chai
- , Mette Aadahl
- , Jie Yao
- , Ani Manichaikul
- , Helen R. Warren
- , Julia Ramirez
- , Jette Bork-Jensen
- , Line L. Kårhus
- , Anuj Goel
- , Maria Sabater-Lleal
- , Raymond Noordam
- , Carlo Sidore
- , Edoardo Fiorillo
- , Aaron F. McDaid
- , Pedro Marques-Vidal
- , Matthias Wielscher
- , Stella Trompet
- , Naveed Sattar
- , Line T. Møllehave
- , Betina H. Thuesen
- , Matthias Munz
- , Lingyao Zeng
- , Jianfeng Huang
- , Bin Yang
- , Alaitz Poveda
- , Azra Kurbasic
- , Claudia Lamina
- , Lukas Forer
- , Markus Scholz
- , Tessel E. Galesloot
- , Jonathan P. Bradfield
- , E. Warwick Daw

- , Joseph M. Zmuda
- , Jonathan S. Mitchell
- , Christian Fuchsberger
- , Henry Christensen
- , Jennifer A. Brody
- , Mary F. Feitosa
- , Mary K. Wojczynski
- , Michael Preuss
- , Massimo Mangino
- , Paraskevi Christofidou
- , Niek Verweij
- , Jan W. Benjamins
- , Jorgen Engmann
- , Rachel L. Kember
- , Roderick C. Slieker
- , Ken Sin Lo
- , Nuno R. Zilhao
- , Phuong Le
- , Marcus E. Kleber
- , Graciela E. Delgado
- , Shaofeng Huo
- , Daisuke D. Ikeda
- , Hiroyuki Iha
- , Jian Yang
- , Jun Liu
- , Hampton L. Leonard
- , Jonathan Marten
- , Börge Schmidt
- , Marina Arendt
- , Laura J. Smyth
- , Marisa Cañadas-Garre
- , Chaolong Wang
- , Masahiro Nakatochi
- , Andrew Wong
- , Nina Hutri-Kähönen
- , Xueling Sim
- , Rui Xia

- , Alicia Huerta-Chagoya
- , Juan Carlos Fernandez-Lopez
- , Valeriya Lyssenko
- , Meraj Ahmed
- , Anne U. Jackson
- , Marguerite R. Irvin
- , Christopher Oldmeadow
- , Han-Na Kim
- , Seungho Ryu
- , Paul R. H. J. Timmers
- , Liubov Arbeeva
- , Rajkumar Dorajoo
- , Leslie A. Lange
- , Xiaoran Chai
- , Gauri Prasad
- , Laura Lorés-Motta
- , Marc Pauper
- , Jirong Long
- , Xiaohui Li
- , Elizabeth Theusch
- , Fumihiko Takeuchi
- , Cassandra N. Spracklen
- , Anu Loukola
- , Sailalitha Bollepalli
- , Sophie C. Warner
- , Ya Xing Wang
- , Wen B. Wei
- , Teresa Nutile
- , Daniela Ruggiero
- , Yun Ju Sung
- , Yi-Jen Hung
- , Shufeng Chen
- , Fangchao Liu
- , Jingyun Yang
- , Katherine A. Kentistou
- , Mathias Gorski
- , Marco Brumat

- , Karina Meidtner
- , Lawrence F. Bielak
- , Jennifer A. Smith
- , Prashantha Hebbar
- , Aliki-Eleni Farmaki
- , Edith Hofer
- , Maoxuan Lin
- , Chao Xue
- , Jifeng Zhang
- , Maria Pina Concas
- , Simona Vaccargiu
- , Peter J. van der Most
- , Niina Pitkänen
- , Brian E. Cade
- , Jiwon Lee
- , Sander W. van der Laan
- , Kumaraswamy Naidu Chitrala
- , Stefan Weiss
- , Martina E. Zimmermann
- , Jong Young Lee
- , Hyeok Sun Choi
- , Maria Nethander
- , Sandra Freitag-Wolf
- , Lorraine Southam
- , Nigel W. Rayner
- , Carol A. Wang
- , Shih-Yi Lin
- , Jun-Sing Wang
- , Christian Couture
- , Leo-Pekka Lyytikäinen
- , Kjell Nikus
- , Gabriel Cuellar-Partida
- , Henrik Vestergaard
- , Bertha Hildalgo
- , Olga Giannakopoulou
- , Qiuyin Cai
- , Morgan O. Obura

- , Jessica van Setten
- , Xiaoyin Li
- , Karen Schwander
- , Natalie Terzikhan
- , Jae Hun Shin
- , Rebecca D. Jackson
- , Alexander P. Reiner
- , Lisa Warsinger Martin
- , Zhengming Chen
- , Liming Li
- , Heather M. Highland
- , Kristin L. Young
- , Takahisa Kawaguchi
- , Joachim Thiery
- , Joshua C. Bis
- , Girish N. Nadkarni
- , Lenore J. Launer
- , Huaixing Li
- , Mike A. Nalls
- , Olli T. Raitakari
- , Sahoko Ichihara
- , Sarah H. Wild
- , Christopher P. Nelson
- , Harry Campbell
- , Susanne Jäger
- , Toru Nabika
- , Fahd Al-Mulla
- , Harri Niinikoski
- , Peter S. Braund
- , Ivana Kolcic
- , Peter Kovacs
- , Tota Giardoglou
- , Tomohiro Katsuya
- , Konain Fatima Bhatti
- , Dominique de Kleijn
- , Gert J. de Borst
- , Eung Kweon Kim

- , Hieab H. H. Adams
- , M. Arfan Ikram
- , Xiaofeng Zhu
- , Folkert W. Asselbergs
- , Adriaan O. Kraaijeveld
- , Joline W. J. Beulens
- , Xiao-Ou Shu
- , Loukianos S. Rallidis
- , Oluf Pedersen
- , Torben Hansen
- , Paul Mitchell
- , Alex W. Hewitt
- , Mika Kähönen
- , Louis Pérusse
- , Claude Bouchard
- , Anke Tönjes
- , Yii-Der Ida Chen
- , Craig E. Pennell
- , Trevor A. Mori
- , Wolfgang Lieb
- , Andre Franke
- , Claes Ohlsson
- , Dan Mellström
- , Yoon Shin Cho
- , Hyejin Lee
- , Jian-Min Yuan
- , Woon-Puay Koh
- , Sang Youl Rhee
- , Jeong-Taek Woo
- , Iris M. Heid
- , Klaus J. Stark
- , Henry Völzke
- , Georg Homuth
- , Michele K. Evans
- , Alan B. Zonderman
- , Ozren Polasek
- , Gerard Pasterkamp

- , Imo E. Hoefer
- , Susan Redline
- , Katja Pahkala
- , Albertine J. Oldehinkel
- , Harold Snieder
- , Ginevra Biino
- , Reinhold Schmidt
- , Helena Schmidt
- , Y. Eugene Chen
- , Stefania Bandinelli
- , George Dedoussis
- , Thangavel Alphonse Thanaraj
- , Sharon L. R. Kardia
- , Norihiro Kato
- , Matthias B. Schulze
- , Giorgia Girotto
- , Bettina Jung
- , Carsten A. Böger
- , Peter K. Joshi
- , David A. Bennett
- , Philip L. De Jager
- , Xiangfeng Lu
- , Vasiliki Mamakou
- , Morris Brown
- , Mark J. Caulfield
- , Patricia B. Munroe
- , Xiuqing Guo
- , Marina Ciullo
- , Jost B. Jonas
- , Nilesh J. Samani
- , Jaakko Kaprio
- , Päivi Pajukanta
- , Linda S. Adair
- , Sonny Augustin Bechayda
- , H. Janaka de Silva
- , Ananda R. Wickremasinghe
- , Ronald M. Krauss

- , Jer-Yuarn Wu
- , Wei Zheng
- , Anneke I. den Hollander
- , Dwaipayan Bharadwaj
- , Adolfo Correa
- , James G. Wilson
- , Lars Lind
- , Chew-Kiat Heng
- , Amanda E. Nelson
- , Yvonne M. Golightly
- , James F. Wilson
- , Brenda Penninx
- , Hyung-Lae Kim
- , John Attia
- , Rodney J. Scott
- , D. C. Rao
- , Donna K. Arnett
- , Mark Walker
- , Heikki A. Koistinen
- , Giriraj R. Chandak
- , Chittaranjan S. Yajnik
- , Josep M. Mercader
- , Teresa Tusié-Luna
- , Carlos A. Aguilar-Salinas
- , Clicerio Gonzalez Villalpando
- , Lorena Orozco
- , Myriam Fornage
- , E. Shyong Tai
- , Rob M. van Dam
- , Terho Lehtimäki
- , Nish Chaturvedi
- , Mitsuhiro Yokota
- , Jianjun Liu
- , Dermot F. Reilly
- , Amy Jayne McKnight
- , Frank Kee
- , Karl-Heinz Jöckel

- , Mark I. McCarthy
- , Colin N. A. Palmer
- , Veronique Vitart
- , Caroline Hayward
- , Eleanor Simonsick
- , Cornelia M. van Duijn
- , Fan Lu
- , Jia Qu
- , Haretsugu Hishigaki
- , Xu Lin
- , Winfried März
- , Esteban J. Parra
- , Miguel Cruz
- , Vilmundur Gudnason
- , Jean-Claude Tardif
- , Guillaume Lettre
- , Leen M. ’t Hart
- , Petra J. M. Elders
- , Scott M. Damrauer
- , Meena Kumari
- , Mika Kivimaki
- , Pim van der Harst
- , Tim D. Spector
- , Ruth J. F. Loos
- , Michael A. Province
- , Bruce M. Psaty
- , Ivan Brandslund
- , Peter P. Pramstaller
- , Kaare Christensen
- , Samuli Ripatti
- , Elisabeth Widén
- , Hakon Hakonarson
- , Struan F. A. Grant
- , Lambertus A. L. M. Kiemeney
- , Jacqueline de Graaf
- , Markus Loeffler
- , Florian Kronenberg

- , Dongfeng Gu
- , Jeanette Erdmann
- , Heribert Schunkert
- , Paul W. Franks
- , Allan Linneberg
- , J. Wouter Jukema
- , Amit V. Khera
- , Minna Männikkö
- , Marjo-Riitta Jarvelin
- , Zoltan Kutalik
- , Francesco Cucca
- , Dennis O. Mook-Kanamori
- , Ko Willems van Dijk
- , Hugh Watkins
- , David P. Strachan
- , Niels Grarup
- , Peter Sever
- , Neil Poulter
- , Jerome I. Rotter
- , Thomas M. Dantoft
- , Fredrik Karpe
- , Matt J. Neville
- , Nicholas J. Timpson
- , Ching-Yu Cheng
- , Tien-Yin Wong
- , Chiea Chuen Khor
- , Charumathi Sabanayagam
- , Annette Peters
- , Christian Gieger
- , Andrew T. Hattersley
- , Nancy L. Pedersen
- , Patrik K. E. Magnusson
- , Dorret I. Boomsma
- , Eco J. C. de Geus
- , L. Adrienne Cupples
- , Joyce B. J. van Meurs
- , Mohsen Ghanbari

- , Penny Gordon-Larsen
- , Wei Huang
- , Young Jin Kim
- , Yasuharu Tabara
- , Nicholas J. Wareham
- , Claudia Langenberg
- , Eleftheria Zeggini
- , Johanna Kuusisto
- , Markku Laakso
- , Erik Ingelsson
- , Goncalo Abecasis
- , John C. Chambers
- , Jaspal S. Kooner
- , Paul S. de Vries
- , Alanna C. Morrison
- , Kari E. North
- , Martha Daviglus
- , Peter Kraft
- , Nicholas G. Martin
- , John B. Whitfield
- , Shahid Abbas
- , Danish Saleheen
- , Robin G. Walters
- , Michael V. Holmes
- , Corri Black
- , Blair H. Smith
- , Anne E. Justice
- , Aris Baras
- , Julie E. Buring
- , Paul M. Ridker
- , Daniel I. Chasman
- , Charles Kooperberg
- , Wei-Qi Wei
- , Gail P. Jarvik
- , Bahram Namjou
- , M. Geoffrey Hayes
- , Marylyn D. Ritchie

- , Pekka Jousilahti
- , Veikko Salomaa
- , Kristian Hveem
- , Bjørn Olav Åsvold
- , Michiaki Kubo
- , Yoichiro Kamatani
- , Yukinori Okada
- , Yoshinori Murakami
- , Unnur Thorsteinsdottir
- , Kari Stefansson
- , Yuk-Lam Ho
- , Julie A. Lynch
- , Daniel J. Rader
- , Philip S. Tsao
- , Kyong-Mi Chang
- , Kelly Cho
- , Christopher J. O'Donnell
- , John M. Gaziano
- , Peter Wilson
- , Charles N. Rotimi
- , Scott Hazelhurst
- , Michèle Ramsay
- , Richard C. Trembath
- , David A. van Heel
- , Gen Tamiya
- , Masayuki Yamamoto
- , Bong-Jo Kim
- , Karen L. Mohlke
- , Timothy M. Frayling
- , Joel N. Hirschhorn
- , Sekar Kathiresan
- , Michael Boehnke
- , Pradeep Natarajan
- , Gina M. Peloso
- , Christopher D. Brown
- , Andrew P. Morris
- , Themistocles L. Assimes

- , Panos Deloukas
- , Yan V. Sun
- & Cristen J. Willer

Corresponding authors

Correspondence to Themistocles L. Assimes or Cristen J. Willer.

Rights and permissions

Reprints and Permissions

About this article

Cite this article

Graham, S.E., Clarke, S.L., Wu, KH.H. *et al.* Author Correction: The power of genetic diversity in genome-wide association studies of lipids. *Nature* **618**, E19–E20 (2023). https://doi.org/10.1038/s41586-023-06194-2

- Published: 26 May 2023

- Issue Date: 15 June 2023

- DOI: https://doi.org/10.1038/s41586-023-06194-2

Share this article

Anyone you share the following link with will be able to read this content:

Get shareable link

Sorry, a shareable link is not currently available for this article.

Copy to clipboard

Provided by the Springer Nature SharedIt content-sharing initiative

This article was downloaded by **calibre** from https://www.nature.com/articles/s41586-023-06194-2

- Author Correction
- Published: 25 May 2023

Author Correction: LACC1 bridges NOS2 and polyamine metabolism in inflammatory macrophages

- Zheng Wei,
- Joonseok Oh,
- Richard A. Flavell &
- …
- Jason M. Crawford

Show authors

Nature **volume 618**, page E21 (2023)

- 522 Accesses
- 1 Altmetric
- Metrics details

The Original Article was published on 17 August 2022

Correction to: *Nature* https://doi.org/10.1038/s41586-022-05111-3
Published online 17 August 2022

In the version of the article originally published, the authors noticed that the unit of time under Methods, ‘Measurement of enzymatic parameters’, for the recombinant LACC1 in vitro activity was inadvertently shown as seconds instead of minutes. Based on this discrepancy, the *x* axes in Extended Data Fig. 2a–e have been corrected to “t(min)”; the *y* axes in Fig. 1c and e, Extended Data Fig. 1e and f and Extended Data Fig. 2f have been corrected to “µM/h”; and in Fig. 1d the units for k_{cat} have been corrected to “h^{-1}” and for k_{cat}/K_m to “h^{-1} M^{-1}.” These measured in vitro parameters could be consistent with reported cyanate values of about 45 nM in human plasma (ref. 23 in the article; Lundquist, P. et al., *Anal. Biochem*. **211**, 23–27 (1993)). The errors have been corrected in the HTML and PDF versions of the article.

Author information

Authors and Affiliations

1. Department of Immunobiology, Yale University School of Medicine, New Haven, CT, USA

   Zheng Wei & Richard A. Flavell

2. Institute of Biomolecular Design & Discovery, Yale University, West Haven, CT, USA

   Zheng Wei, Joonseok Oh & Jason M. Crawford

3. Department of Chemistry, Yale University, New Haven, CT, USA

   Joonseok Oh & Jason M. Crawford

4. Howard Hughes Medical Institute, Yale University School of Medicine, New Haven, CT, USA

   Richard A. Flavell

5. Department of Microbial Pathogenesis, Yale University School of Medicine, New Haven, CT, USA

Jason M. Crawford

Corresponding authors

Correspondence to Richard A. Flavell or Jason M. Crawford.

Rights and permissions

Reprints and Permissions

About this article

Cite this article

Wei, Z., Oh, J., Flavell, R.A. *et al.* Author Correction: LACC1 bridges NOS2 and polyamine metabolism in inflammatory macrophages. *Nature* **618**, E21 (2023). https://doi.org/10.1038/s41586-023-06244-9

- Published: 25 May 2023
- Issue Date: 15 June 2023
- DOI: https://doi.org/10.1038/s41586-023-06244-9

Share this article

Anyone you share the following link with will be able to read this content:

Get shareable link

Sorry, a shareable link is not currently available for this article.

Copy to clipboard

Provided by the Springer Nature SharedIt content-sharing initiative

---

This article was downloaded by **calibre** from https://www.nature.com/articles/s41586-023-06244-9

| Section menu | Main menu |

- Author Correction
- Published: 26 May 2023

Author Correction: Testosterone histories from tusks reveal woolly mammoth musth episodes

- Michael D. Cherney ,
- Daniel C. Fisher,
- Richard J. Auchus,
- Adam N. Rountrey,
- Perrin Selcer,
- Ethan A. Shirley,
- Scott G. Beld,
- Bernard Buigues,
- Dick Mol,
- Gennady G. Boeskorov,
- Sergey L. Vartanyan &
- …
- Alexei N. Tikhonov

Show authors

Nature **volume 618**, page E22 (2023)

- 401 Accesses

- 2 Altmetric

- Metrics details

The Original Article was published on 03 May 2023

Correction to: *Nature* https://doi.org/10.1038/s41586-023-06020-9
Published online 3 May 2023

In the version of the article originally published, there was a typographical error in the "Female woolly mammoth" subsection, where the landmass of Wrangel Island was described as 76,000 km^2 rather than 7,600 km^2; the error has been corrected in the HTML and PDF versions of the article.

Author information

Authors and Affiliations

1. Museum of Paleontology, University of Michigan, Ann Arbor, MI, USA

   Michael D. Cherney, Daniel C. Fisher, Adam N. Rountrey, Ethan A. Shirley & Scott G. Beld

2. Medical School, University of Michigan, Ann Arbor, MI, USA

   Michael D. Cherney & Richard J. Auchus

3. Earth and Environmental Sciences, University of Michigan, Ann Arbor, MI, USA

   Daniel C. Fisher & Ethan A. Shirley

4. Program in the Environment, University of Michigan, Ann Arbor, MI, USA

   Perrin Selcer

5. Mammuthus, St Jean de Luz, France

   Bernard Buigues

6. Natural History Museum, Rotterdam, Netherlands

   Dick Mol

7. Diamond and Precious Metals Geology Institute, Siberian Branch of the Russian Academy of Sciences, St Jean de Luz, Russian Federation

   Gennady G. Boeskorov

8. Laboratory of Cenozoic Geology and Paleogeography, North-east Interdisciplinary Scientific Research Institute, Far-East Branch, Russian Academy of Sciences, Magadan, Russian Federation

   Sergey L. Vartanyan

9. Laboratory of Mammals, Zoological Institute of the Russian Academy of Sciences, St Petersburg, Russian Federation

   Alexei N. Tikhonov

10. Mammoth Museum, North-Eastern Federal University, Yakutsk, Russian Federation

    Alexei N. Tikhonov

Corresponding author

Correspondence to Michael D. Cherney.

Rights and permissions

Reprints and Permissions

About this article

Cite this article

Cherney, M.D., Fisher, D.C., Auchus, R.J. *et al.* Author Correction: Testosterone histories from tusks reveal woolly mammoth musth episodes. *Nature* **618**, E22 (2023). https://doi.org/10.1038/s41586-023-06242-x

- Published: 26 May 2023
- Issue Date: 15 June 2023
- DOI: https://doi.org/10.1038/s41586-023-06242-x

Share this article

Anyone you share the following link with will be able to read this content:

Get shareable link

Sorry, a shareable link is not currently available for this article.

Copy to clipboard

Provided by the Springer Nature SharedIt content-sharing initiative

---

This article was downloaded by **calibre** from https://www.nature.com/articles/s41586-023-06242-x

- Author Correction
- Published: 31 May 2023

Author Correction: Short tRNA anticodon stem and mutant eRF1 allow stop codon reassignment

- Ambar Kachale,
- Zuzana Pavlíková,
- Anna Nenarokova,
- Adriana Roithová,
- Ignacio M. Durante,
- Petra Miletínová,
- Kristína Záhonová,
- Serafim Nenarokov,
- Jan Votýpka,
- Eva Horáková,
- Robert L. Ross,
- Vyacheslav Yurchenko,
- Petra Beznosková,
- Zdeněk Paris ,
- Leoš Shivaya Valášek &
- …
- Julius Lukeš

Show authors

Nature **volume 618**, page E23 (2023)

- 410 Accesses

- 3 Altmetric

- Metrics details

The Original Article was published on 11 January 2023

Correction to: *Nature* https://doi.org/10.1038/s41586-022-05584-2
Published online 11 January 2023

In the version of this article originally published, there was an error in the northern blot of control $tRNA^{Glu}_{UUC}$ in the lower “aminoacylation” panel of Figure 2b, where the wrong blot–the same as for $tRNA^{Glu}_{CUC}$–was used by mistake. The correct blots showing charging of $tRNA^{Glu}_{UUC}$ and $tRNA^{Glu}_{CUC}$ controls were included in Supplementary Figure 1, showing the uncropped gel source data. This erroneous blot in Figure 2b has now been replaced with the correct one from Supplementary Figure 1. This oversight does not affect the conclusions of this article. The change is made in the HTML and PDF versions of the article.

Author information

Author notes

1. Anna Nenarokova

   Present address: School of Biological Sciences, University of Bristol, Bristol, UK

2. Eva Horáková

   Present address: Institute of Microbiology, Czech Academy of Sciences, Třeboň, Czech Republic

3. These authors contributed equally: Ambar Kachale, Zuzana Pavlíková, Anna Nenarokova

Authors and Affiliations

1. Institute of Parasitology, Biology Centre, Czech Academy of Sciences, České Budějovice, Czech Republic

   Ambar Kachale, Anna Nenarokova, Ignacio M. Durante, Kristína Záhonová, Serafim Nenarokov, Jan Votýpka, Eva Horáková, Zdeněk Paris & Julius Lukeš

2. Faculty of Sciences, University of South Bohemia, České Budějovice, Czech Republic

   Ambar Kachale, Anna Nenarokova, Serafim Nenarokov, Zdeněk Paris & Julius Lukeš

3. Institute of Microbiology, Czech Academy of Sciences, Prague, Czech Republic

   Zuzana Pavlíková, Adriana Roithová, Petra Miletínová, Petra Beznosková & Leoš Shivaya Valášek

4. Faculty of Science, Charles University, BIOCEV, Prague, Czech Republic

   Kristína Záhonová & Jan Votýpka

5. Life Science Research Centre, Faculty of Science, University of Ostrava, Ostrava, Czech Republic

   Kristína Záhonová & Vyacheslav Yurchenko

6. Thermo Fisher Scientific, Franklin, MA, USA

   Robert L. Ross

Corresponding authors

Correspondence to Zdeněk Paris, Leoš Shivaya Valášek or Julius Lukeš.

Rights and permissions

Reprints and Permissions

About this article

Cite this article

Kachale, A., Pavlíková, Z., Nenarokova, A. *et al.* Author Correction: Short tRNA anticodon stem and mutant eRF1 allow stop codon reassignment. *Nature* **618**, E23 (2023). https://doi.org/10.1038/s41586-023-06260-9

- Published: 31 May 2023
- Issue Date: 15 June 2023
- DOI: https://doi.org/10.1038/s41586-023-06260-9

Share this article

Anyone you share the following link with will be able to read this content:

Get shareable link

Sorry, a shareable link is not currently available for this article.

Copy to clipboard

Provided by the Springer Nature SharedIt content-sharing initiative

This article was downloaded by **calibre** from https://www.nature.com/articles/s41586-023-06260-9

- Author Correction
- Published: 31 May 2023

Author Correction: Quantifying the cost savings of global solar photovoltaic supply chains

- John Paul Helveston,
- Gang He &
- Michael R. Davidson

Nature **volume 618**, page E24 (2023)

- 357 Accesses
- 2 Altmetric
- Metrics details

The Original Article was published on 26 October 2022

Correction to: *Nature* https://doi.org/10.1038/s41586-022-05316-6
Published online 26 October 2022

In the version of this article initially published, the projected 2030 solar PV capacity of the United States in the sustainable development scenario in Table 1, now reading as 411 GW, mistakenly stated 628 GW. Furthermore, the implied CAGR (%), now shown as 16, was originally written as 21. These revisions changed the projected 2030 prices in Fig. 4 for the United States under the sustainable development scenario in the national market (orange line) from US$276 per kW (25% higher than with global market) to

US$287 per kW (30% higher than with global market). In addition, the cumulative savings of the United States from 2020 to 2030 in the sustainable development scenario in Extended Data Fig. 1 has been corrected to US$12 (11–13) billion from US$18 (17–20) billion. Text related to these data have also been updated to reflect the changes. In the second-to-last sentence of the Abstract, "20–25 per cent" now reads as "20–30 per cent". In the sentence beginning "Under the sustainable development…" in the second paragraph of the 'Future trajectories' subsection, "25%" has been replaced with "25–30%" and "US$276.2 per kW versus US$221.3 per kW in the United States" now reads as "US$287.2 per kW versus US$221.3 per kW in the United States". In the last sentence of this paragraph, "US$36 (33–39)" has been updated to "US$29 (27–32)".

The projected global market size is directly cited from the IEA World Energy Outlook 2020 Sustainable Development Scenario projection, not summarized from independent country sizes, and is therefore not affected by the proposed revisions. The errors have been corrected in the HTML and PDF versions of the article.

The authors thank Yiting Xu and Jing Meng at University College London for identifying this data point error. For transparency, the differences between the original and revised versions of Table 1, Fig. 4, Extended Data Fig. 1, and text from the Abstract and 'Future trajectories' subsection are shown in the Supplementary Information attached to this notice.

Author information

Authors and Affiliations

1. Department of Engineering Management and Systems Engineering, George Washington University, Washington DC, USA

   John Paul Helveston

2. Department of Technology and Society, College of Engineering and Applied Sciences, Stony Brook University, Stony Brook, NY, USA

Gang He

3. Department of Mechanical and Aerospace Engineering, University of California San Diego, La Jolla, CA, USA

   Michael R. Davidson

4. School of Global Policy and Strategy, University of California San Diego, La Jolla, CA, USA

   Michael R. Davidson

Corresponding author

Correspondence to Gang He.

Supplementary information

Supplementary Information

Rights and permissions

Reprints and Permissions

About this article

Cite this article

Helveston, J.P., He, G. & Davidson, M.R. Author Correction: Quantifying the cost savings of global solar photovoltaic supply chains. *Nature* **618**, E24 (2023). https://doi.org/10.1038/s41586-023-06262-7

- Published: 31 May 2023
- Issue Date: 15 June 2023

- DOI: https://doi.org/10.1038/s41586-023-06262-7

Share this article

Anyone you share the following link with will be able to read this content:

Get shareable link

Sorry, a shareable link is not currently available for this article.

Copy to clipboard

Provided by the Springer Nature SharedIt content-sharing initiative

---

This article was downloaded by **calibre** from https://www.nature.com/articles/s41586-023-06262-7

Spotlight

- **Nutrition research in China** [14 Jun 2023]
  • China today is perhaps unrecognisable from the same country 50 years ago. And as the country has changed, so has its diet. How will these deep shifts in the country’s eating habits affect healthcare and treatment, and what can scientists and policymakers do about it?



Why China's changing diet is a bellyache for public health

The shift from traditional cuisine to an increasingly Western diet is creating issues for nutrition researchers and countering other steps forward in public health.

- Yvaine Ye
- Jack Leeming

Spotlight14 Jun 2023 Nature



From tea to tofu: why Chinese dietary staples are rich pickings for research

Public-health researchers are exploring the nutritional benefits of some familiar favourites of the Chinese table.

- Michael Eisenstein

Spotlight14 Jun 2023 Nature

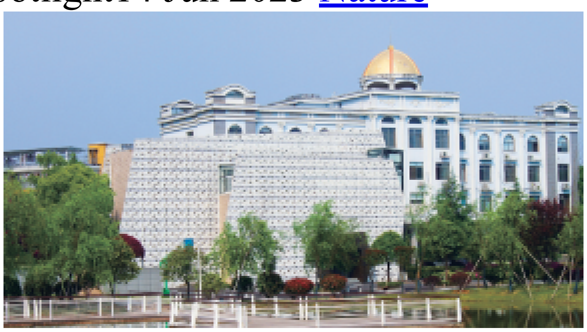


Mapping the Chinese microbiome: it’s time for a united effort

Much has been done to investigate the human microbiome. Now, it’s time for countries to dig deeper.

- Ju-Sheng Zheng

Spotlight14 Jun 2023 Nature



Peeling back the science of citrus fruits

A conversation with Yang Shan, a member of Chinese Academy of Engineering

Advertisement Feature15 Jun 2023



The science of healthier pigs

A conversation with Yulong Yin, a member of the Chinese Academy of Engineering

Advertisement Feature15 Jun 2023



Weighing up the evidence for multivitamins

A large-scale study in china, spanning more than three decades, has explored the effect of multivitamins to treat dietary deficiencies.

Advertisement Feature14 Jun 2023



Exploring new frontiers in health management through nutrition research

A nutrition company in china is forging high-profile collaborations with researchers to understand the mechanisms behind disease and the ageing process.

Advertisement Feature14 Jun 2023



Massive library of bacterial strains could lead to menu of new fermented foods

Researchers in china have established a library of more than 6,000 strains of lactic acid bacteria, which could help fuel a new industry of innovative fermented food products.

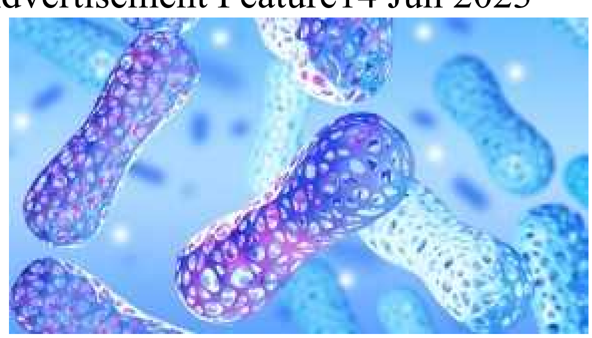
Advertisement Feature14 Jun 2023



Unravelling the mysteries of Mānuka honey

Detailed research into the unique make-up of New Zealand's popular Mānuka honey is marching into a new chapter.

Advertisement Feature14 Jun 2023



Exploring links between probiotics and reducing obesity

In a pre-clinical study, researchers at a university in China and a dairy manufacturer explore mechanisms that could connect obesity reduction and probiotics.

Advertisement Feature14 Jun 2023

This article was downloaded by **calibre** from https://www.nature.com/collections/fjgjggijdh